Risk-attitudes and Socio-economic Characteristics of Economic Migrants: An Empirical Analysis of U.S. Interstate Migration

Inaugural – Dissertation

zur

Erlangung der wirtschaftswissenschaftlichen Doktorwürde

des Fachbereichs Wirtschaftswissenschaften

der Philipps-Universität Marburg

eingereicht von:

Sophia Koch

Dipl.-Volkswirtin aus Gießen

Erstgutachter: Prof. Dr. Bernhard Nietert

Zweitgutachter: Prof. Dr. Bernd Hayo

Einreichungsdatum: 6. Januar 2014

Prüfungstermin: 12. Februar 2014

Erscheinungsort: Marburg

Hochschulkennziffer: 1180

Table of Contents

List of Tables

List of Figures

List of Abbreviations

e.g.	for example (abbreviation for *exempli gratia*)
etc.	and so on (abbreviation for *et cetera*)
f.	and following page (abbreviation for *folio*)
ff.	and following pages (abbreviation for *foliis*)
i.e.	that is (abbreviation for *id est*)
p.	page
pp.	pages
ACS	American Community Survey
Ad1	Index indicating risk-attitudes based on income parameters estimated from income data of three consecutive years around the year of the move (actual year of the move, the year preceding, and the one following the move) without inflation adjustment
Ad2	Index indicating risk-attitudes based on income parameters estimated from income data of three consecutive years around the year of the move (actual year of the move, the year preceding, and the one following the move) with inflation adjustment by the annual U.S. Consumer Price Index
Ann	Index indicating risk-attitudes based on an annual planning period
Ed1	Index indicating risk-attitudes based on education definition one
Ed2	Index indicating risk-attitudes based on education definition two
Ed3	Index indicating risk-attitudes based on education definition three
Ed4	Index indicating risk-attitudes based on education definition four
Ed5	Index indicating risk-attitudes based on education definition five
Ed6	Index indicating risk-attitudes based on education definition six
Fam	Index indicating risk-attitudes based on family income
Ind	Index indicating risk-attitudes based on Head’s individual income
Lif	Index indicating risk-attitudes based on a planning period until reaching life expectancy
LP2	Index indicating risk-attitudes based on risk-measure semi-variance (i.e., Lower Partial Moment Two)

L1	Index indicating risk-attitudes based on minimizing predictive errors as defined by L_1-norm (i.e., sum of predictive errors)
L2	Index indicating risk-attitudes based on minimizing predictive errors as defined by L_2-norm (i.e., sum of squared predictive errors)
Ma	Index indicating risk-attitudes based on minimizing predictive errors as defined by L_∞-norm (i.e., maximum predictive error)
One	Index indicating risk-attitudes based on income parameters estimated from annual income data in the year of the move, i.e., data of only one year
Poo	Index indicating risk-attitudes based on income parameters that are estimated from samples where people are clustered into groups by a pooled clustering over all years, i.e., the cluster definitions applied are the same in all years of data
PSID	Panel Study of Income Dynamics
Sep	Index indicating risk-attitudes based on income parameters that are estimated from samples where people are clustered into groups separately for each year of data, i.e., the cluster definitions applied are the different in each year of data
Unw	Index indicating risk-attitudes based on income parameters that are estimated from unweighted samples
Var	Index indicating risk-attitudes based on risk-measure variance
Wei	Index indicating risk-attitudes based on income parameters that are estimated from weighted samples representative of the U.S. population
Wor	Index indicating risk-attitudes based on a planning period until reaching full retirement age (working life)
2Kat1	Transformation rule that transforms the numerical results on risk-attitudes to a binary variable by dividing all people into risk-averse and risk-seeking migrants applying a threshold value of zero and deleting the 5% of migrants from the sample that are least pronounced in their degree of risk-attitude (i.e., 5% of migrants that ar4e closest to being risk-neutral)
2Kat3	Transformation rule that transforms the numerical results on risk-attitudes to a binary variable by dividing all people into risk-averse and risk-seeking migrants applying a threshold value of zero
4Kat2	Transformation rule that transforms the numerical results on risk-attitudes to a binary variable by dividing all people into risk-averse and risk-seeking migrants applying a threshold value of zero and restricting the sample to those migrants that belong (i) either to the most extreme half among the risk-averse or (ii) the most extreme half among the risk-seeking migrants (i.e., only 50% of the migrants in my sample are left for the analysis)

List of Symbols

c	Number of variables to be entered in sets in each regression
d	destination state actually chosen
i	Index for decision-makers
$inc_{i,j}$	Total annual income net of migration costs of decision-maker i in destination state j with planning period one year
$inc_{i,j,t,T}$	Present value of total income net of migration costs of decision-maker i in destination state j at time t concerning planning period $T-t$
j	Index for destination states $j = 1, \dots, N$
$m_{x_{test}}$	Index for regression run on test variable x_{test} where $m_{x_{test}} = 1, \dots, M_{x_{test}}$
p	Positive integer with $p = 1,2, \dots, \infty$
q_{t,T_i}	Risk-free discount factor that discounts cash-flows at time T_i to time t
rnf	Number of the remaining non-fixed variables that are not the test variable x_{test}
r_{t,θ_i}	Risk-free spot rate per annum for capital tie-up $\theta_i - t$
t	Time index for the time the migration decision is taken
$var_t(.)$	Variance operator at the time the migration decision is taken t
$w_{m_{x_{test}}}$	Weight put on regression $m_{x_{test}}$ testing test variable x_{test}
$\boldsymbol{x}_{fixed}$	Vector of fixed variables in the extreme bounds analysis
$\boldsymbol{x}_{non-fixed}$	Vector of non-fixed variables in the extreme bounds analysis
x_{test}	Test variable in the extreme bounds analysis
$CDF(0)$	Maximum of (i) the area under the density function of the normal distribution lying below zero and (ii) the area under the density function of the normal distribution lying above zero
$E\{.\}$	Unconditional expected value operator
$E_t(.)$	Expected value operator at time t
$H0_{Binomial}$	Null hypothesis tested by the binomial test
$H0_{t\text{-}test}$	Null hypothesis tested by the student t-test

$\mathrm{L}_{m_{x_{test}}}$	Likelihood of regression $m_{x_{test}}$ testing test variable x_{test}
N	Number of possible destination states $j = 1, \dots, N$
P	Random variable
T	Time index for the time when the planning period ends
T_i	Time index for the time when the planning period of decision-maker i ends where $\mathrm{T} = \{\mathrm{t}+1,\ \text{time when full retirement age is reached, time when life expectancy is reached}\}$
$T - t$	Planning period for which the decision to migrate is taken
$U(.)$	Utility function operator
W	Standard normally distributed random variable
X	Random variable
Y	Normally distributed random variable
Z	Log-normally distributed random variable
α	Risk-attitude parameter
$\alpha_{i,t,T}$	Risk-attitude of decision-maker i at the time the migration decision is taken t concerning a planning period of $T - t$
$\boldsymbol{\beta}_{x_{fixed}, m_{x_{test}}}$	Vector of regression coefficients of the fixed variables resulting from regression $m_{x_{test}}$ on test variable x_{test}
$\boldsymbol{\beta}_{x_{non-fixed}, m_{x_{test}}}$	Vector of regression coefficients of the non-fixed variables resulting from regression $m_{x_{test}}$ on test variable x_{test}
$\beta_{x_{test}, m_{x_{test}}}$	Regression coefficient of test variable x_{test} resulting from regression $m_{x_{test}}$ on test variable x_{test}
$\beta_{0, m_{x_{test}}}$	Intercept of regression $m_{x_{test}}$ on test variable x_{test}
$\overline{\beta_{x_{test}}}$	Average regression coefficient of test variable x_{test}
γ	Dependent variable in the extreme bounds analysis
η	Index for U.S. residents with the same socio-economic characteristics like the decision-maker where $\eta = 1, \dots, n$
θ_i	Time index for decision-maker i with $\theta_i = t+1, \dots, T_i$
λ	Probability of being risk-averse
μ	Mean
σ	Standard deviation

$\sigma_{x_{test},m_{x_{test}}}$	Robust standard error of the regression coefficient of the test variable x_{test} resulting from regression $m_{x_{test}}$ on test variable x_{test}
$\overline{\sigma_{x_{test}}{}^2}$	Average of squared robust standard errors of test variable x_{test}
$\upsilon_{i,j,t,T}(.)$	Predictive error operator of the theoretical migration decision model concerning the actual migration decision of individual i, for destination state j at time t and planning period $T-t$
$\psi(.)$	Preference value operator
$\psi_{i,j,t,T}(.)$	Preference value operator of decision-maker i for destination state j at time t and planning period $T-t$
$\boldsymbol{\Upsilon}_{i,t,T}(.)$	Vector of predictive errors of the theoretical migration decision model concerning the actual migration decision of individual i at time t and planning period $T-t$

Part A – Migration Decision Model

1 Introduction

1.1 The role of risk-attitudes in determining migration decisions

The World Bank estimates that in 2010 approximately 214 million people lived outside their country of birth, 92% of them moved for economic reasons.[1] Economic migration is always associated with a considerable amount of risk due to a new environment, a depreciation of skills in the destination region, or the danger of being unemployed.[2] Given the great number of migrants and their tremendous effect on the economies of the sending and receiving countries, it is important to understand the migration decision in the face of uncertainty. This requires, in particular, that the risk-attitudes of migrants when taking the migration decision is characterized.

This characterization of migrants' risk-attitudes has one immediate practical application that translates into an additional research question. In the ongoing debate on labor migration in industrialized countries it has often been argued that young and well-educated migrants are needed in order to reduce the demographic pressure and the shortage of skilled labor. This raises the question of how the desired group of migrants can be attracted. Since migration is associated with risk, this question can only be answered if the relation between socio-economic characteristics and risk-attitudes in the migration context is known.

1.2 Gap in the migration literature regarding risk-attitudes

Migration literature can be divided into three categories. The first strand of literature relates to theoretical and empirical studies that fully neglect risk from the migrant's perspective. Instead, these

[1] World Bank (2011), p. 268. The remaining 8% are refuges or asylum seekers. The numbers are in line with those reported by the United Nations online data tool (see United Nations, Department of Economic and Social Affairs, Population Division (2009)). Numbers refer to the year 2010.

[2] See for example, Todaro (1969), p. 140, Harris and Todaro (1970), p. 127, Smith (1979), p. 31, Stark and Levhari (1982), p. 191 f., Stark (1984), p. 207, Burda (1995), p. 4 f., Rotte and Vogler (1999), p. 4, Bosco (2000), p. 3, Chen, Chiang, and Leung (2001), p.1, Locher (2001), p.2, Epstein (2002), p. 6, Epstein and Gang (2002), p. 1, Khwaja (2002), p. 3, Mahmood and Schömann (2003), p.4, Anam, Chiang, and Hua (2008), p. 238, Ruangsiri (2004), p. 2, Constant and Zimmermann (2005), p. 2, Moretto and Vergalli (2005), p. 2, Vergalli (2006), p. 5, Krupka and Donaldson (2008), p. 7, Demiralp (2009), p. 12, d'Haultfoeuille and Maurel (2009), p. 2, Gibson and McKenzie (2009), p. 21, Umblijs (2012), p. 4.

papers either argue under certainty[3] or focus on uncertainty regarding the quality of migrants from the perspective of the receiving countries[4].

The second strand of literature relates to theoretical papers that incorporate risk in their migration decision models. Unfortunately, the great majority of these papers simplify the analysis by assuming either risk-neutral[5] or risk-averse migrants[6] without providing any proof regarding the assumed risk-attitude.[7] These simplifications might result in misleading predictions concerning migration behavior for the following reasons: First, there is no consensus concerning the risk-attitude of migrants. While it has often been argued that migrants are risk-seeking or at least less risk-averse than their non-moving counterparts,[8] other authors find the opposite effect in their empirical studies.[9] Second, it has been shown that people are very heterogeneous concerning their attitudes towards risk.[10] Consequently, they cannot be considered to be homogeneously risk-neutral or risk-averse. The only theoretical papers on migration that explicitly account for all types of risk-attitudes (i.e., risk-aversion, risk-neutrality, and risk-seeking) are those of Ruangsiri (2004) and Heitmüller (2005). Unfortunately, both migration models are not applied to real world data. Hence, they cannot contribute to the characterization of migrants' true risk-attitudes in the migration context.

[3] See for example, Stark (1984) in his second and third migration model. Empirical papers that fully neglect risk are, for example, Peridy (2006) and Liebig and Sousa-Poza (2004).

[4] See literature on self-selection going back to the Roy-Model (1951) such as Borjas (1987, 1994, 1999a), Borjas, Bronars, and Trejo (1992), Bauer (2002), Sheldon (2002), Brücker and Trübswetter (2004), Brücker and Defoort (2006), Peridy (2006). And most contributions concerned with welfare migration like Wildasin (1991, 1994), Sinn and Werding (2001), Sinn (2004), Sinn and Ochel (2004).

[5] Migration models assuming risk-neutral migrants are, for example, Todaro (1969), Harris and Todaro (1970), Burda (1995), Giannetti (1999), Bosco (2000), Davies, Greenwood, and Li (2001), Locher (2001), Epstein (2002), Mahmood and Schömann (2003), Pedersen, Pytlikova, and Smith (2004), Vergallo and Moretto (2005), Dostie and Léger (2006), Vergalli (2006), Kennan and Walker (2008), Demiralp (2009), d'Haultfoeuille and Maurel (2009), Mitze and Reibowski (2010), Dequiedt and Zenou (2011).

[6] Migration models assuming risk-averse migrants are, for example, Smith (1979), Stark and Levhari (1982), Stark (1984) in his first migration decision model of the paper, Dustmann (1996, 2001), Daveri and Faini (1999), Chen, Chiang, and Leung (2001), Halliday (2008), Otrachshenko and Popova (2012).

[7] Some authors analyze the migration decision separately for assumed risk-neutrality and risk-aversion like Stark (1981), Khwaja (2002), Anam, Chiang, and Hua (2008), Bayer and Juessen (2008), Guler, Guvenen, and Violante (2010).

[8] For example, Katz and Stark (1986) argue that one explanation for people moving in the presence of a negative income differential is that migrants are risk-seeking (see Katz and Stark (1986), p. 135). This perspective is supported by Heitmüller (2005). Using a utility framework for the migration decision under risk, he shows in his theoretical model that people having the lowest level of risk-aversion (usually those being risk-seeking) are the first to migrate (see Heitmüller (2005), p. 8). Other authors find empirical evidence that a higher level of self-reported willingness to take risk raises the probability of migrating (see for example, Kan (2003), p. 585, Jaeger, Dohmen, Falk, Huffman, Sunde, and Bonin (2008), p. 7, or Gibson and McKenzie (2011), p. 25).

[9] For example, Bonin, Constant, Tatsiramos, and Zimmermann (2006) find that migrants are more risk-averse than natives (see Bonin, Constant, Tatsiramos, and Zimmermann (2006), p. 4).

[10] See for example, Barsky, Juster, Komball, and Shapiro (1997), p. 545, Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009), p. 8, Ding, Hartog, and Sun (2010), p. 10.

The third strand of literature relates to empirical papers that introduce risk-attitudes as an explanatory variable to the migration decision.[11] For example, Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005) and Jaeger, Dohmen, Falk, Huffman, Sunde, Bonin (2008).[12] Both papers use data of the 2004 wave of the German Socio-Economic Panel (SOEP) to explain the probability of migrating from Eastern and Western Germany by self-assessed willingness to take risks. While Jaeger, Dohmen, Falk, Huffman, Sunde, Bonin (2008) only apply general willingness to take risks, Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005) additionally use several non-migration domain-specific willingnesses to take risks. Both approaches are afflicted with three problems: First, they do not contribute to the characterization of risk-attitudes but use it as an explanatory variable. Second, risk-attitudes are not based on actual choices but on hypothetical questions. Since risk-attitudes derived from choices actually taken have shown to be most predictive for future behavior, results from hypothetically asked questions about willingness to take risks might be misleading.[13] Third, risk-attitudes have been proven to be domain-specific[14], i.e., the risk-attitude of one and the same person usually varies with context.[15] For example, a person can be risk-averse concerning investment decisions but risk-seeking when doing sports. It is therefore necessary to restrict the analysis to the domain of migration if migration behavior is of interest, rather than employing general risk-attitudes.

1.3 Research questions

Given the relevance of economic migration and the identified gaps in the migration literature, I formulate three research questions: First, what are migrants' individual risk-attitudes in the context of economic migration? Second, does risk actually plays a significant role in the economic migration decision, i.e., are economic migrants significantly different from being risk-neutral. Third, how are economic migrants' risk-attitudes related to their socio-economic characteristics such as gender, age, and education?

---

[11] Studies belonging to the third strand of literature must be distinguished from those studies that simply compare self-assessed willingness to take risk of the native and migrant population like Bonin, Constant, Tatsiramos, and Zimmermann (2006).
[12] An earlier version of this paper is Jaeger, Bonin, Dohmen, Falk, Huffman, and Sunde (2007).
[13] Holt and Laury (2002) compare behavior in a hypothetical lottery and a real-payoff lottery and conclude that behaviors are very different (see Holt and Laury (2002), p. 1654).
[14] See for example, empirical findings of Weber, Blais, and Betz (2002), p. 282, Blais and Weber (2006), p. 41, Ding, Hartog, and Sun (2010), p. 5, Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005), p. 33. Nicholson, Soane, Fenton-O'Creevy, and Willman (2005) also find domain-specific risk-attitudes, but not for all individuals with some being strictly risk-seeking or risk-avoiding in all domains (see Nicholson, Soane, Fenton-O'Creevy, and Willman (2005), p. 170).
[15] For an empirical proof based on a large and representative sample see Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005), p. 30, and Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009), p. 23.

1.4 Methodology

All research questions will be addressed empirically. Since I am interested in economic migration risk, non-economic determinants influencing the migration decision must be eliminated. To be able to isolate the effect of economic risk, U.S. interstate migration is considered for the empirical analysis. Within the U.S., other sources of risk like lingual and cultural barriers, differences in the political systems etc. do not exert a distorting influence and, hence, can be neglected. Investigating U.S. interstate migration, the following methodology is applied: To answer the first research question of my dissertation, migrants' risk-attitudes are estimated by calibrating a mean-variance decision model to the true migration decision among 321 U.S. interstate migrants. The calibration is performed by non-linear minimization of predictive errors that are measured by L_p-norms one, two, and infinity. The second research question, whether risk actually plays a significant role in the migration decision, is answered by applying two alternative statistical tests, namely a simple mean difference test and a binomial test. For the third research question of my dissertation, the analysis of the relation of socio-economic characteristics and risk-attitudes in the migration context, an extreme bounds analysis using binary logistic regression is run.

1.5 Results

The empirical results of my dissertation can be briefly summarized as follows: Concerning the first research question, I find that the 321 migrants of my sample include both risk-averse and risk-seeking migrants. Consequently, economic migrants cannot be considered to be homogeneously risk-neutral or risk-averse.

Concerning the second research question, I find that, first, risk actually plays a significant role in the economic migration decision and second, migrants are significantly risk-averse.

Concerning the third research question, I conclude that the relation between socio-economic characteristics and risk-attitudes of economic migrants crucially depends on the way risk-attitudes are estimated.

1.6 Contribution to the literature

In view of the existing studies on migration, my dissertation contributes to the literature in four ways: First, my theoretical model is the first migration decision model that accounts for all types of risk-attitudes (risk-aversion, risk-neutrality, and risk-seeking) and is applied to real data at the same time. To the best of my knowledge, only Heitmüller (2002, 2005) and Ruangsiri (2004) have developed a migration decision framework under which all types of risk-attitudes were possible. Yet, both models have not been applied to real data. Second, I am the first to empirically estimate risk-attitudes in the context of migration. In comparison to the empirical literature that relates risk-attitudes to the migration decision like Jaeger, Dohmen, Falk, Huffman, Sunde, and Bonin (2008) and Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005) my approach has two advantages: it is empirically estimated based on real migration choices rather than only a self-assessed willingness to take risks, and it is directly related to the context of economic migration rather than the general willingness to take risks in other non-migration domains such as finance, career, sports or leisure. Third, my empirical analysis is the first to accurately differentiate between risk-averse, risk-neutral, and risk-seeking individuals in the context of migration rather than giving only a tendency of willingness to take risk. The migration decision model of Heitmüller (2002, 2005) is potentially able to do the same thing, however, he does not perform an empirical analysis. Instead, he arbitrarily chooses values for variables of his model to derive tendencies of migration behavior. Therefore, he does not contribute to the characterization of migrants' risk-attitudes. Fourth, socio-economic characteristics on migration-specific risk-attitudes are investigated for the first time. This is due to the fact that risk-attitudes of migrants have not been estimated in the migration domain so far.

1.7 Organization of the dissertation

The remainder of this dissertation is organized as follows: In Part A a migration decision model under risk that accounts for individual risk-attitudes is developed. Part B documents the amount of data cleaning that is necessary to create a data set which is suitable for the empirical analysis. Part C estimates individual risk-attitudes, tests whether risk plays a significant role in the migration decision, and relates risk-attitudes to socio-economic characteristics of migrants.

2 Theoretical migration decision model under risk

To model the economic migration decision, I develop a decision model that accounts for all types of risk-attitudes (i.e., risk-aversion, risk-neutrality, and risk-seeking). The model reads as follows: A person has to choose one out of many possible destination locations in order to improve his income. To formalize the idea of income improvement, a more detailed analysis of the decision problem is required, i.e., the decision-maker, the decision field, and the objective function must be specified.

2.1 Decision-maker

Generally, migration decisions might affect several people. For that reason, it is by no means clear who takes the migration decision.

In my model, the decision-maker is assumed to be a single person. His decision can either cover only himself moving on his own or a group of people moving together as a family. In the latter case the decision-maker is taking the decision on behalf of the whole family. In other words, no collective decision-making as in social choice theory is modeled.

2.2 Decision field

The decision field comprises the time of decision-making, the planning period, the set of alternatives, the outcome of the decision, and the states of nature.

2.2.1 Time of decision-making

This analysis considers a static model, i.e., the migration decision is taken only once. Moreover, the point in time at which the migration-decision is taken is assumed to be identical to the moving date (i.e., when the migration actually takes place). The intuition behind the correspondence of migration decision date and moving date is as follows: The decision-maker seeks to obtain the best information in the sense of the latest information available. Hence, he decides about where to move not before the decision actually must be taken, i.e., the time of the move.

2.2.2 Planning period

As a static one-period model is considered in this study, the planning period is defined to be the next period only. Nevertheless, the length of this single period can be arbitrarily chosen. While a period of one year seems plausible at first sight, a longer planning period seems reasonable. First, individuals usually do not decide about migration every single year but rather stay for several years – sometimes until they die. To account for the different perspectives, I consider three different planning periods: (i) the period of one year; (ii) the period until reaching full retirement age; (iii) the period until reaching life expectancy.

2.2.3 Set of alternatives: geographical locations

The set of alternatives the decision-maker faces are the possible destination locations, including the null alternative of staying in the home location.

2.2.4 Outcome: total income net of migration costs

Outcomes are consequences of migration decisions. Since this analysis is restricted to economically motivated migration, outcome is defined as total income net of migration costs. Generally, non-monetary outcome determinants are also possible. They are neglected, however, in this analysis because for economically motivated migration it is reasonable to assume that the driving force of migration can be expressed in terms of total income net of migration costs.

According to Borjas (1999b) migration costs include three types of costs: (i) direct costs for transportation of persons and belongings, (ii) opportunity costs of foregone earnings during migration, and (iii) psychic costs, resulting from displeasure felt from leaving the familiar environment of family, friends and social networks.[16] Additionally, a fourth type of migration cost is often mentioned in the literature: costs arising from differences in culture and language barriers since this may affect the transferability of human capital.[17] Except for psychic costs, migration costs can be at least approximately measured in monetary terms and thus be incorporated in the income parameters in the sense of total income net of migration costs.

In general, for the migration decision both (i) decision-maker's individual income and (ii) total family income could be relevant. The latter definition is due to the existence of so called tied-movers, i.e., people that "participate in moves that result in a net loss for themselves but positive net returns for the family".[18] Net loss means, an alternative is chosen that is not optimally from an individual's, but from the family's perspective. For example, assume the decision-maker earns 1,000 U.S. Dollars in location A and 1,500 U.S. Dollars in location B, while the second person moving with him earns

[16] See Borjas (1999b), p. 1711 for the first three types of migration costs and Sjaastad (1962), p. 85 for psychic costs of migration.
[17] See for example, Mayda (2005), p. 13.
[18] Compton and Pollak (2004), p. 6. The same contents can also be found in Mincer (1978), p. 751.

2,000 U.S. Dollars in location A and 1,300 U.S. Dollars in location B. If the decision-maker decides based on his own income, he would decide to move to location B where he earns most. In contrast, if the migration decision was taken based on family income, location A would have been chosen because family income is higher there (3,000 U.S. Dollars compared to 2,800 U.S. Dollars). In this example, the decision-maker is the tied mover because for him a net loss of 500 U.S. Dollars occurs. Neglecting the existence of tied movers would mean that each decision-maker only considers his own income. In contrast, family income must be considered to account for the existence of tied movers. In order to account for both perspectives, the decision problem is analyzed for individual and family income separately.

Total income on the family and individual level includes all types of income, i.e., earned income, welfare income, retirement income and income from capital investments.

2.2.5 Decision under risk

The migration decision is risky if and only if (i) deciding for one and the same alternative can result in more than one outcome depending on the state of nature where (ii) the probability of each state is known.

In my context, the state of nature is the job someone gets, which includes unemployment as one potential state of nature. While income is certain for each job, it is never certain which job and, therefore, income someone gets in any location. Hence, income is stochastic. To better understand this argument, consider two scenarios. In the first scenario, the decision-maker moves without having a job in the new location. Clearly, the migrant cannot be sure about which job he gets or whether he will get any job at all. Income will be stochastic. In the second scenario, the decision-maker has a job acceptance for the destination location before the move takes place. In this case it can still happen that he loses his job and falls into welfare. Of course, it is also possible that he gets promoted to a better job. Consequently, here again, the outcome of the migration decision (i.e., income) is stochastic.

Since income for a given job is certain, the probability distribution for income and jobs coincide for the decision-maker.

2.3 Objectives and their weighting

It is assumed that the decision-maker pursues two goals when taking the migration decision: The first goal is concerned with maximizing the opportunities of migration. This can be expressed by expected value of total income net of migration costs. The second goal is concerned with the variations in total income net of migration costs, i.e., the risk aspect. It can be captured by a risk-measure, here its variance.

These two objectives can be conflicting, neutral, or complementary depending on the individual's risk-attitude. While all people try to maximize their expected value a ceteris paribus higher variance is only appreciated by risk-seeking individuals. Risk-averse people try to minimize variance for any given level of expected value. Therefore, a parameter on risk-attitude is introduced to weight the two goals.

2.4 Formal decision model

2.4.1 General model

A model that is principally able to formalize mean (μ) -variance (σ^2) trade off is the so-called hybrid model, which can be formalized in the general form of

$$\psi(\mu, \sigma) = \mu - \alpha \cdot \sigma^2 \tag{1}$$

where $\psi(.)$ denotes the preference value, μ mean, σ^2 variance, and α risk-preference parameter.

Since the actual risk-attitude of the decision-maker is not known yet, the model incorporates all types of risk-attitude: A risk-averse person would have a positive value of α, while a risk-seeker would have a negative value in the same variable. Risk-neutrality can be modeled by assuming $\alpha = 0$.[19]

The risk-parameter is specific to decision-maker i (or the respective family), the time the decision is taken t, and the planning period $(T - t)$. The reasons for this are as follows: First, it seems reasonable to assume that risk-attitudes depend on the socio-economic characteristics of the decision-maker. Second, risk-attitude is time variable because one and the same individual might have a changing risk-attitude as he turns older and gains more experience. Third, the risk involved in

[19] For a proof that the preference function of Equation (1) is indeed capable of dealing with risk-averse and risk-seeking individuals please refer to the Appendix, p. 365.

a lifetime decision may be different compared to the decision about moving to another place only for the next year. Risk-attitude is therefore attributed to the planning period.

The same logic holds for income parameters mean and variance. They are specific to decision-maker i (or the respective family), the location j, the time of decision-making t, and the length of the planning period $(T-t)$. The rationale behind these specifications is as follows: First, mean and variance of total income net of migration costs depend on socio-economic characteristics of the decision-maker (or the respective family members of the migrating family; note that family income results from the sum of personal income over all family members). For example, a well-educated individual usually earns more than a less educated individual all other things being equal, while the variance of income might be higher for the latter. To account for all socio-economic characteristics of the decision-maker (or the respective family members) that affect income parameters, income must be indexed i. Second, it could be that one and the same person faces different income parameters in different locations, for example, due to a different relation of labor demand and supply for individuals with these very socio-economic characteristics. Third, income parameters might vary over years due to the current economic situation. Forth, income parameters depend on the length of the time period for which the migration decision is taken. Clearly mean and variance of annual income are not identical to mean and variance of the present value of lifetime income. To account for different planning periods, T is specified to be either next year, the time when full retirement age is reached, or the time when life expectancy is reached.

Given these specifications, the general hybrid decision model of Equation (1) can be formalized as follows

$$\max_j \psi_{i,j,t,T}\left(inc_{i,j,t,T}\right) = \max_j \left\{E_t\left(inc_{i,j,t,T}\right) - \alpha_{i,t,T} var_t\left(inc_{i,j,t,T}\right)\right\}. \qquad (2)$$

where $j = 0,..,N$ denotes the different locations with $j = 0$ being the home location, i the decision-maker, t the time index for the time the migration decision is taken, T the time index for the time the planning period ends with $\mathrm{T} \in$ {t+1, time when full retirement age is reached, time when life expectancy is reached}, $inc_{i,j,t,T}$ the present value of total income net of migration costs of decision-maker i in location j at time t concerning planning period $(T-t)$, $E_t(.)$ the expected value operator at time t, and $var_t(.)$ the variance operator at time t.[20]

[20] Note that the preference value of the optimal destination location does not necessarily have to be positive to trigger migration. Imagine an individual with a negative risk-adjusted expected income in the home location $\left(\psi_{\mathrm{i,j=0,t,T}} < 0\right)$. For such a person moving would pay-off as long as the preference function of the chosen destination location is less negative than the one of the home location.

Verbally, Equation (2) describes the preference value $\psi_{i,j,t,T}$ decision-maker i maximizes by choosing one out of many possible destination locations $j = 0,..,N$ at the time of the migration decision t considering income parameters over the planning period $(T - t)$.

2.4.2 Estimating income parameters for different planning periods

The true income parameters mean and variance of decision-makers and their respective families cannot be directly observed but must be estimated. Because family income is the sum of family members' individual income, the following section is concerned with the estimation of individual's income parameters only. Arguments and derivations can be analogously applied to families' income parameters.

The income parameters $E_t(inc_{i,j,\mathrm{t},T})$ and $var_t(inc_{i,j,\mathrm{t},T})$ of individual i with a certain set of socio-economic characteristics can be estimated based on income parameters observable at the time of the move t among all residents of the respective location j that share the same socio-economic characteristics. This approach can be chosen because the income of people having the same socio-economic characteristics can be interpreted as realizations of the random variable job of that specific individual. Concerning the planning period $(T - t)$ for which income must be estimated, three different time periods are considered in my study.

Usually income is measured in terms of annual income $(T = next\ year = t + 1)$ because a period of one year includes not only monthly payments but also additional payments like Christmas allowance and other bonus payments. Statistics on annual income from which mean and variance can be estimated are publicly available as cross-section data. For longer planning periods (working life and lifetime) the available annual income parameters must be transformed to a period of more than a year. The approach chosen in this analysis argues as follows exemplary for lifetime income, but can be applied to working life income analogously:

Lifetime income

Decision-maker i considers his time until reaching his life expectancy T_i not as a random variable but as a constant which equals the data given in the mortality table. This is consistent with decision-makers being risk-neutral concerning mortality risk. Furthermore, the decision-maker expects his annual income given a certain job in a certain location j to stay constant over his remaining lifetime so that for each job it holds

$$inc_{i,j,t,t+1} = inc_{i,j,t+1,t+2} = \cdots = inc_{i,j,T_i-1,T_i} \equiv inc_{i,j} \tag{3}$$

where t denotes the year of the move, $t+1$ the next year, T_i the time until decision-maker i reaches his life expectancy as reported in the mortality table, and $inc_{i,j}$ the constant total annual income net of migration costs of decision-maker i in destination state j.

Lifetime income as present value of future cash-flows

Consequently, lifetime income of individual i at time t equals the present value of the cash-flow from income over his remaining lifetime. Analogously to the planning period of one year where a state of nature is defined as having a certain job, a state of nature for longer planning periods is having a certain job over several years. This means for each job (i.e., state of nature) lifetime income is deterministic (see Equation (3)) so that the present value of lifetime income can simply be computed by discounting the income stream by means of the risk-free term structure at the time of the move t:

$$inc_{i,j,t,T_i} = \sum_{\theta_i=t+1}^{T_i} \frac{inc_{i,j}}{e^{\theta_i \cdot r_{t,\theta_i}}} = inc_{i,j} \sum_{\theta_i=t+1}^{T_i} \frac{1}{e^{\theta_i \cdot r_{t,\theta_i}}} \tag{4}$$

where θ_i denotes the time index with $\theta_i = t+1, \ldots, T_i$ and r_{t,θ_i} the risk-free spot rate per annum for capital tie-up $\theta_i - t$.

Estimation of mean and variance of lifetime income

Here again, the related income parameters mean and variance can be estimated based on observable realizations of lifetime income among $\eta = 1, \ldots, n$ residents of location j having the same socio-economic characteristics like the migrant at time t.

Since income parameters are estimated by a sample of people belonging to the same cohort as the migrant, they have the same life expectancy so that

$$T_{\eta=1} = T_{\eta=2} = \cdots = T_{\eta=n} = T_i \tag{5}$$

where $\eta = 1, \ldots, n$ denotes the index for residents with the same socio-economic characteristics like the decision-maker.

The estimation of expected value of lifetime income from the sample can be written as mean of all lifetime incomes observable

$$E_t\left(inc_{i,j,t,T_i}\right) = \frac{\sum_{\eta=1}^{n} inc_{\eta,j,t,T_i}}{n}. \tag{6}$$

Given that lifetime income can be written as risk-free discounted annual income (Equation (4)) and life expectancy is constant within the sample (Equation (5)), the expected value of lifetime income reads

$$E_t\left(inc_{i,j,t,T_i}\right) = \frac{\sum_{\eta=1}^{n}\left(inc_{\eta,j}\sum_{\theta=t+1}^{T_i}\frac{1}{e^{\theta\cdot r_{t,\theta}}}\right)}{n} \tag{7}$$

which can further be simplified to

$$E_t\left(inc_{i,j,t,T_i}\right) = \frac{\left(\sum_{\eta=1}^{n} inc_{\eta,j}\right)\left(\sum_{\theta=t+1}^{T_i}\frac{1}{e^{\theta\cdot r_{t,\theta}}}\right)}{n} = E_t\left(inc_{i,j}\right)\underbrace{\left(\sum_{\theta=t+1}^{T_i}\frac{1}{e^{\theta\cdot r_{t,\theta}}}\right)}_{q_{t,T_i}} \tag{8}$$

where q_{t,T_i} denotes the risk-free discount factor that discounts cash-flows at time T_i to the time the migration decision is taken t.

Obviously the expected value of lifetime income can be derived from the expected value of annual income $E\left(inc_{i,j}\right)$ multiplied by a risk-free discount factor q_{t,T_i}.

The second parameter to be estimated from the sample of $\eta = 1, \dots, n$ residents of location j is variance of lifetime income. The sample estimator can be written as

$$var_t\left(inc_{i,j,t,T_i}\right) = \frac{1}{n-1}\sum_{\eta=1}^{n}\left[inc_{\eta,j,t,T_i} - E_t\left(inc_{i,j,t,T_i}\right)\right]^2. \tag{9}$$

Using Equation (4) (lifetime income equals the present value of lifetime cash-flows from annual income) and Equation (8), Equation (9) can be written as

$$var_t\left(inc_{i,j,t,T_i}\right)$$

$$= \frac{1}{n-1}\sum_{\eta=1}^{n}\left[inc_{\eta,j}\cdot\underbrace{\sum_{\theta=t+1}^{T_i}\frac{1}{e^{\theta\cdot r_{t,\theta}}}}_{q_{t,T_i}} - E_t\left(inc_{i,j}\right)\underbrace{\left(\sum_{\theta=t+1}^{T_i}\frac{1}{e^{\theta\cdot r_{t,\theta}}}\right)}_{q_{t,T_i}}\right]^2 \tag{10}$$

and further simplified to

$$var_t(inc_{i,j,t,T_i}) = {q_{t,T_i}}^2 \cdot \underbrace{\frac{1}{n-1}\sum_{\eta=1}^{n}\left[inc_{i,j} - E_t(inc_{i,j})\right]^2}_{var_t(inc_{i,j})}. \tag{11}$$

The lifetime income parameters of individual i given in Equations (8) and (11) can thus be derived from the annual income parameters $\left(E_t(inc_{i,j}), var_t(inc_{i,j})\right)$ observable among the n residents of location j having the same socio-economic characteristics like individual i by multiplying the annual estimators with the risk-free discount factor q_{t,T_i} specific to the decision-maker's planning period $(T_i - t)$.

3 Empirical analysis

3.1 Empirical application: U.S. interstate migration

To isolate the effect of economic risk, the theoretical migration decision model is applied to U.S. interstate migration. Within the U.S., other sources of risk like language and cultural barriers, differences in the political systems etc. do not exert a distorting influence and, hence, can be neglected.

For the empirical analysis, an object of investigation must be found that comes closest to the assumptions of the decision model. The crucial assumption, which makes the analysis difficult, is the measurement of migration costs in monetary terms. This assumption relates to direct costs for transportation of persons and belongings, opportunity costs of foregone earnings during migration, and costs due to reduced transferability of human capital arising from cultural differences and language barriers.

These assumptions and the focus on monetary determinants in the migration decision model result in two requirements for the perfect object for the empirical analysis: First, the income parameters in different locations must be different in order to trigger migration. Second, different locations should be identical concerning all other characteristics in order to keep migration costs constant for all locations. If this was the case, the incorporation of migration costs in the income parameters can be neglected without distortions because they would not alter the rank of the preference values assigned to a certain location.

In reality, such conditions cannot be found. While locations would be relatively similar when a small area is considered, they may not possess very different income possibilities. When greater territories such as continents are considered, locations may have greater differences in their income parameters, but differences in other characteristics such as culture or language will also be greater.

To solve this trade-off U.S. interstate migration is considered. This results in several advantages: First, income parameters of U.S. states differ quite a lot (see Figure 11, p. 109). Second, direct migration costs for transportation of persons and belongings are composed of fixed and variable costs. Fixed costs are, for example, costs for packing, loading and unloading of the household (eventually with help of professionals), the one-time rent for a moving truck, furniture blankets, a doll and further equipment. They can be assumed to be comparable across U.S. states. Variable costs include costs for gas and wages for eventually hired professionals driving the truck. They get smaller in relation to fixed costs the smaller the distance of the move is. Restricting the analysis to U.S. interstate migration also restricts the maximum distance between states, thus restricting the maximum amount of variable costs in relation to fixed costs. If it is assumed that direct migration costs are at large fixed costs, which are comparable between U.S. states, direct migration costs can be neglected in case of U.S. interstate migration. Furthermore, both variable costs and fixed costs are relatively small in comparison to total working life income or total lifetime income. Third, within the U.S. the education system is comparable, differences in culture and language are relatively small compared to international migration, e.g., from a rural Mexican area to New York City. It can therefore be assumed that migration costs due to reduced transferability of human capital are negligible.

3.2 Risk-measures: variance versus semi-variance

The theoretical migration decision model derived in Part A, Section 2 is able to account for all types of risk-attitudes (i.e., risk-aversion, risk-neutrality, and risk-seeking). This is possible because the risk-measure variance contains both positive and negative deviations from the expected value. That way the needs of risk-averse decision-makers are captured by the negative deviation whereas those of risk-seeking decision-makers are captured by the positive deviations. However, integrating positive deviations might not adequately capture the risk for a risk-averse decision maker. For that reason, an alternative risk-measure is considered as well, namely semi-variance. Semi-variance accounts only for the downside-risk in that it considers solely negative deviation from the expected income.[21]

Since my dissertation is concerned with an empirical analysis of economic migration where it is not clear which risk-measure is actually applied by migrants' when taking the migration decision, I apply semi-variance as an alternative risk-measure in my empirical analysis.

3.3 Data requirements and data creation

In order to empirically analyze the migration decision based on the migration decision model derived above, a suitable data set is needed. Unfortunately, no single data set exists that meets the requirements for the empirical analysis of U.S. interstate migration. To illustrate the complexity of finding a suitable data set, the data requirements are outlined below:

1) It must be possible to identify U.S. interstate migrants and
2) to restrict the analysis to economically motivated migration. This means, the reason to move must be indicated.
3) Family members must be clearly defined in order to estimate family income.

[21] Note that semi-variance is equivalent to Lower Partial Moment Two with reference value expected value.

4) To control for other determinants that may interfere with pure economic reasons to migrate, family ties of the decision-maker before and after the move must be available. For example, it could be that someone reports economic reasons for his move but at the same time he has decided to leave his old partner and children. In this case it could be that the family composition change interferes with the reported (economic) reason to move and even may contradict it. This would bias the risk-attitude derived from the migration decision observed. In order to account for these potentially interfering constellations, family ties before and after the move must be available.
5) The data must include individual's socio-economic characteristics gender, age, and education.
6) It must be possible to estimate the income parameters these migrants face in each possible destination state depending on their socio-economic characteristics.

There are data sets which partially meet these requirements, but unfortunately, a single data set meeting all requirements does not exist.

The only solution possible is to create a new data set by merging various data sets. In case of U.S. interstate migration, two data sets are needed:

1) The Panel Study of Income Dynamics (PSID) meets requirements one to five (i.e., it is possible to identify US-interstate migrants, the reason to move is included, and each individual's socio-economic characteristics and family ties are included). Although income is reported for each person, the sample size is not large enough to estimate individual's income parameters in each U.S. state depending on gender, age, and education. Furthermore, data from 1999 to 2009 is only available every second year.
2) The American Community Survey (ACS), which the U.S. Census Bureau recommends for income studies, contains detailed demographic characteristics on the subnational level like U.S. states.[22] It has a much higher sample size, is available for every year from 1999 to 2009, and meets all above mentioned requirements with only one exception: the reason to move (requirement two).

Hence, it must be put special emphasis on obtaining the required data set. This can be done by merging two types of data from two different sources. First, migrants with their family ties, socio-economic characteristics (gender, age, and education), and their reason to move can be identified with help of the Panel Study of Income Dynamics (PSID). Second, the income parameters that these

[22] See U. S. Census Bureau (2010).

migrants face in each possible destination state depending on their socio-economic characteristics can be estimated from the American Community Survey (ACS). Third, the data sets from the first two steps must be merged in order to obtain the data set including all the above mentioned requirements.

All steps are afflicted with multiple and complex problems that result from (i) different data structures of the different data sets, (ii) changing data structures over years of data within the same source and (iii) missing information that must be brought forward from earlier waves of data.

These problems can be solved by separate data cleaning of both types of data sets, namely the Panel Study of Income Dynamics and the American Community Survey. The cleaning of both data sets and the merging is documented and discussed in detail in Part B. The next section presents both data sets and its key characteristics important to understand the need for data cleaning, and they provide key definitions of my dissertation.

3.4 Description of data sets

3.4.1 Panel Study of Income Dynamics

3.4.1.1 Objective and history

The initial wave of this household panel was conducted in 1968 as a representative longitudinal sample of U.S. individuals.[23] The study is run under the supervision of the Survey Research Center at the Institute for Social Research (University of Michigan) and is sponsored by government agencies, foundations, and other organizations.[24] The central focus of the survey is on income, employment, family composition changes, and demographic events.[25] Over the years the sample size has grown from almost 5,000 families in 1968 to roughly 8,900 families and 71,000 individuals in 2009.[26] This was possible because of high wave-to-wave response rates of 96%-98%[27] and the success in following family split-offs, i.e., when young adults establish their own economic family unit or other family members leaving the household.[28]

---

[23] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 21, Paragraph 1.
[24] See Institute for Social Research, Survey Research Center, University of Michigan (2011e).
[25] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 21, Paragraph 1.
[26] See Institute for Social Research, Survey Research Center, University of Michigan [No date b], p.2.
[27] See Institute for Social Research, Survey Research Center, University of Michigan [No date a], p. 6.
[28] See Institute for Social Research, Survey Research Center, University of Michigan [No date b], p.2 and Panel Study of Income Dynamics, public use dataset (2011).

3.4.1.2 File structure[29]

The Panel Study of Income Dynamics consists of two types of data files: (i) several Single-Year Family Files giving annual cross-sectional information on families and (ii) one Cross-Year Individual File including panel data on individuals.[30] Both data sets of the Panel Study of Income Dynamics are needed and must be merged to gain all information needed for my analysis as the following sections show. In order to understand the need for data cleaning and the problems confronted with when data sets are merged, Figure 1, p. 20 gives the stylized file structure of these two types of data sets, exemplary for waves 2007 and 2009. Additionally, Table 1, p. 21 gives an overview of the variables included in both data sets. The following subsections explain how to read Figure 1, p. 20 and Table 1, p. 21.

[29] See Institute for Social Research, Survey Research Center, University of Michigan (2011d).
[30] See Institute for Social Research, Survey Research Center, University of Michigan (2011d), Paragraph A.

Single-Year Family File 2007

Family identifier

Head's background information

Family interview number 07	Residence 07	…	Age of Head	Sex of Head	…	Educated in U.S.	Finished high school	…	Highest college degree	…
1		…			…			…		…
2		…			…			…		…
…		…			…			…		…
8,289		…			…			…		…

Single-Year Family File 2009

Family identifier

Head's background information

Family interview number 09	Residence 09	…	Age of Head	Sex of Head	…	Educated in U.S.	Finished high school	…	Highest college degree	…
1		…			…			…		…
2		…			…			…		…
…		…			…			…		…
8,690		…			…			…		…

Cross-Year Individual File 1968-2009

Individual data 1968 … Individual data 2007 Individual data 2009

Family interview number 1968	Person number	Sequence Number `68	…	Family interview number 2007	Sequence Number 2007	…	Family interview number 2009	Sequence Number 2009	…
1	1	1	…	13	1	…	3,547	1	…
1	30	2	…	245	1	…	78	2	…
2	1	1	…	4,698	1	…	5,000	81	…
2	4	2		4,698	2		5,000	1	
2	16	3	…	3	1	…	6,976	1	…
…	…	…	…	…	…	…	…	…	…

Figure 1: Stylized structure of Single-Year Family Files and the Cross-Year Individual File.
Source: Own diagram based on Panel Study of Income Dynamics, public use dataset (2011).

Single-Year Family File 2007		Single-Year Family File 2009	
HOUSING			
ER36002	2007 FAMILY INTERVIEW (ID) NUMBER	ER42001	2009 FAMILY INTERVIEW (ID) NUMBER
ER36003	PSID STATE OF RESIDENCE	ER42002	PSID STATE OF RESIDENCE
…	…	…	…
ER36016	NUMBER OF PEOPLE IN FAMILY UNIT	ER42001	NUMBER OF PEOPLE IN FAMILY UNIT
ER36017	AGE OF HEAD	ER42002	AGE OF HEAD
ER36018	SEX OF HEAD	ER42001	SEX OF HEAD
ER36104	MONTH MOVED	ER42133	MONTH MOVED
ER36105	YEAR MOVED	ER42134	YEAR MOVED
ER36106	WHY MOVED	ER42135	WHY MOVED
…	…	…	…
BACKGROUND AND EDUCATION OF HEAD			
ER40527	WHETHER NEW HEAD IN FAMILY UNIT	ER46504	WHETHER NEW HEAD IN FAMILY UNIT
ER40573	WHETHER HEAD EDUCATED IN US	ER46551	WHETHER HEAD EDUCATED IN US
ER40574	WHETHER GRADUATED HIGHSCHOOL-HEAD	ER46552	WHETHER GRADUATED HIGHSCHOOL-HEAD
ER40582	GRADE OF SCHOOL FINISHED-HEAD	ER46560	GRADE OF SCHOOL FINISHED-HEAD
ER40585	WHETHER ATTENDED COLLEGE-HEAD	ER46563	WHETHER ATTENDED COLLEGE-HEAD
ER40588	HIGHEST YEAR COLLEGE COMPLETED-HEAD	ER46566	HIGHEST YEAR COLLEGE COMPLETED-HEAD
ER40589	WHETHER RECORDED COLLEGE DEGREE-HEAD	ER46567	WHETHER RECORDED COLLEGE DEGREE-HEAD
ER40590	HIGHEST COLLEGE DEGREE RECORDED-HEAD	ER46568	HIGHEST COLLEGE DEGREE RECORDED-HEAD
ER40593	YEARS FOREIGN EDUCATION-HEAD	ER46571	YEARS FOREIGN EDUCATION-HEAD
ER40594	FOREIGN DEGREES-HEAD	ER46572	FOREIGN DEGREES-HEAD
…	…	…	…
BACKGROUND AND EDUCATION OF WIFE			
ER40438	WHETHER NEW WIFE IN FAMILY UNIT	ER46410	WHETHER NEW WIFE IN FAMILY UNIT
	… same variables like Head …		… same variables like Head …
ER40501	FOREIGN DEGREES-WIFE	ER46478	FOREIGN DEGREES-WIFE
…	…	…	…
⇨ **5,069 variables on 8,289 families**		⇨ **5,012 variables on 8,690 families**	

Cross-Year Individual File 1968-2009		
1968		
	ER30000	RELEASE NUMBER
	ER30001	1968 INTERVIEW NUMBER
	ER30002	PERSON NUMBER 68
	ER30003	RELATIONSHIP TO HEAD 68
	ER30004	AGE OF INDIVIDUAL 68
	ER30005	MARR PAIRS INDICATOR 68
	ER30009	IN SCHOOL 68
	ER30010	YRS SCHL COMPL 68
	…	…
1969 - … - 2007		
	…	…
	…	…
2009		
	ER34001	2009 INTERVIEW NUMBER
	ER34002	SEQUENCE NUMBER 09
	ER34003	RELATION TO HEAD 09
	ER34004	AGE OF INDIVIDUAL 09
	ER34005	MONTH INDIVIDUAL BORN 09
	ER34006	YEAR INDIVIDUAL BORN 09
	ER34007	MARITAL PAIRS INDICATOR 09
	…	…
	ER34016	EMPLOYMENT STATUS 09
	ER34017	MONTH LAST IN SCHOOL 09
	ER34018	YEAR LAST IN SCHOOL 09
	ER34019	WHETHER STUDENT 09
	ER34020	YEARS COMPLETED EDUCATION 09
	…	…
⇨ **1,446 variables on 71,285 individuals**		

Table 1: Stylized structure of Single-Year Family Files and the Cross-Year Individual File.
Source: Own diagram based on the Panel Study of Income Dynamics, public use dataset (2011).

3.4.1.2.1 Single-Year Family Files

Single-Year Family Files are released as cross-sectional data for each year of the study from 1968 to 2009. "Each Single-Year Family File contains one record for each family interviewed in the specified year (…) [and, insertion by author] contains all of the family-level variables collected in that wave. The records in each file are identified by the Family Interview Numbers of that year"[31] (see Figure 1, p. 20 and Table 1, p. 21).

As can be seen in Figure 1, p. 20 the Single-Year Family File contains data of 8,289 families in 2007 and 8,690 families in 2009. Each row in the dataset stands for one single family. It contains information on the family level such as their current state of residence, the number of people residing in the family, their housing situation, their income and other family level variables.

3.4.1.2.2 Cross-Year Individual File and Sequence Numbers[32]

The Cross-Year Individual File is released as panel data and contains one record for each person ever in the Panel Study of Income Dynamics. It includes all individual-level variables from 1968 through 2009. The file also contains Family Interview Numbers of the family with which the person was associated in each wave. The history of Family Interview Numbers for each person makes it possible to identify one and the same person over the different waves of Single-Year Family Files. The Cross-Year Individual File also includes the Sequence Number of that person within his current family. The **Sequence Number** ranks people according to their position within the family hierarchy.[33] The head of a family is always numbered 1, the second person, usually the wife, is numbered 2 and so on. In 2009 the Cross-Year Individual File contained 71,285 individuals and their personal data (see Table 1, p. 21).

3.4.1.2.3 Family interview numbers

It is important to note that **Family Interview Numbers** identify families within a certain wave of data, but are time inconsistent. This means, they most certainly change from year to year for one and the same family.[34] The reason for this is that annual interview numbers are assigned based on receipts of the interview; the first interview coming in from field is numbered 1, the second 2, and so on.[35] This means, if family data from the Single-Year Family Files of consecutive waves is merged by Family

---

[31] See Institute for Social Research, Survey Research Center, University of Michigan (2011d), Paragraph A.1.
[32] See Institute for Social Research, Survey Research Center, University of Michigan (2011d), Paragraph A.2.
[33] See Institute for Social Research, Survey Research Center, University of Michigan (2011d), Paragraph C.
[34] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3.
[35] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3.

Interview Numbers, its time inconsistency results in merging different families. For example, family X may have a Family Interview Number of 10 in 2001, while the same Family Interview Number may be given to family Y in 2003. If Single-Year Family Files of 2001 and 2003 would be merged by the Family Interview Number, family X and family Y would be falsely considered as one and the same family.

To account for the time inconsistency of the Family Interview Numbers, it is therefore necessary to make a detour when merging family data: First, data from the Single-Year Family File and the Cross-Year Individual File can be merged by the Family Interview Number of the respective year. Second, when Single-Year Family Files of different waves (i.e., years) must be merged, Family Interview Numbers of all waves must be available. They can be found in the Cross-Year Individual File.

3.4.1.2.4 Background information

In the Panel Study of Income Dynamics "background information" refers to all variables on Head's and Wife's personal background. It is of great relevance to this study because it also includes very detailed information on socio-economic characteristics of Heads and Wives. Unfortunately, questions on socio-economic characteristics are not asked again every wave, but only once when Head/Wife first became Head/Wife of that family.[36] Families that have the same Head/Wife as last wave will not be asked again and background information variables have missing values. For that reason, data must be brought forward from previous waves.

3.4.2 American Community Survey

3.4.2.1 Objective and history

The American Community Survey (ACS) is a nationwide high-density household survey conducted by the United States Census Bureau to replace the U.S. Decennial Census long-form questionnaire on an annual rather than decennial basis.[37] It is designed to provide communities with up-to date data on demographics, education, income, housing etc. of the U.S. population[38] in order to help legal authorities to plan investments and services and decide on the spending of more than $400 billion in federal and state funds each year.[39] Therefore, data is collected continuously over the year nearly every day and aggregated over a specific time period (1, 3, and 5 years).[40]

[36] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 85.
[37] See U.S. Census Bureau (2009c), p. 2-1 and U.S. Bureau of Labor Statistics (2011), Question 1.
[38] See U.S. Census Bureau (2008), p.1.
[39] See U.S. Census Bureau (2012a).
[40] See U.S. Census Bureau (2009a), p. A-1.

While the first test period of the American Community Survey started in 1994 including only a few cities, it was only in 2000 when the U.S. Census Bureau carried out a large-scale, nationwide survey that can be used for nationwide research.[41] From the year 2000 on, the sample size was successively increased till the American Community Survey was finally fully implemented in 2005 including 2.5% of the U.S. population.[42]

3.4.2.2 Data availability: Integrated Public Use Microdata Sample

The raw and full sample of the American Community Sample is not available to the public.[43] Among the publicly available Public Use Microdata Samples (PUMS) of the American Community Survey, the most comfortable on, namely the Integrated Public Use Microdata Sample Series (IPUMS) created under the supervision of the Minnesota Population Research Center at the University of Minnesota, is chosen for the analysis. The advantage of IPUMS files is that it "assigns uniform codes across all the samples and brings relevant documentation into a coherent form".[44] This spares most effort usually needed to clean the data of several waves, such as standardizing possible answers and their coding. In contrast to the full American Community Survey which capture 2.5% of U.S. population, the IPUMS ACS files only cover 1% of the population.[45] Still, the IPUMS ACS files are high-density samples with around 3 million respondents each year.[46] Furthermore, all data sets are available as weighted samples where personal and household weights can be applied to perform analysis representative on the national level as well as geographically smaller subareas. Therefore, the IPUMS ACS files rather than any other U.S. household survey are recommended by the U.S. Census Bureau for cross-section estimates of income on the state level by detailed demographic characteristics.[47] As this is exactly what is needed for my analysis, the IPUMS ACS files are the first choice type of data for my dissertation.

[41] See U.S. Census Bureau (2009c), p. 2-1.
[42] See U.S. Census Bureau (2009c), pp. 2-1 and 12-6.
[43] See U.S. Census Bureau (2012b).
[44] Minnesota Population Center, University of Minnesota [No date a].
[45] See U.S. Census Bureau (2009a), p. 2.
[46] See Integrated Public Use Microdata Sample Series files of the American Community Survey by Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek (2011).
[47] See U.S. Census Bureau (2010), p.1.

4 Key definitions: family, decision-maker, and migrant

In the theoretical migration decision model it was assumed that a decision-maker exists who decides on behalf of the whole family based on (i) his own income or (ii) family income. So far, no clear definition of family and decision-maker exists. Because individual migrants as well as migrating families are identified in the Panel Study of Income Dynamics, the definitions in my dissertation follow the definitions applied in the Panel Study of Income Dynamics. Part A, Section 4.1 defines family, Part A, Section 4.2 defines the decision-maker within the family ("head of family") and Part A, Section 4.3 defines the resulting definition of migrants applied for this analysis.

4.1 Definition of family – Family Unit versus Household Unit

Within the Panel Study of Income Dynamics there is no definition of "family". Instead the survey concentrates on family units (FU) since families are made up of individuals and the composition of those individuals may change from wave to wave. The family unit is defined "as a group of people living together as a family. They are almost always related by blood, marriage, or adoption"[48], but unrelated persons can be part of a family unit if they are "permanently living together and share both income and expenses".[49] Thus, an individual living on his own can form his own family unit with one person only. As the Panel Study of Income Dynamics is about family unit members only,[50] my analysis also concentrates on family units.

To avoid a misunderstanding of family unit, the household unit (HU) is defined as the physical boundary, such as a house or apartment, where members of the Panel Study of Income Dynamics family units reside.[51] Note, that "not everyone living in a household unit is automatically part of the family unit."[52] Household units are irrelevant to my study.

The study at hand follows the convention of the Panel Study of Income Dynamics where "family" always refers to the family unit and "family member" always refers to members of the family unit.

[48] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 80, Paragraph 1.
[49] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 80, Paragraph 2.
[50] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 80, Paragraph 5.
[51] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 80, Paragraph 4.
[52] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 80, Paragraph 5.

4.2 The decision-maker: head of family

Of all family members the head of family (simply called Head) is of particular importance for my study. He is the representative of the whole family unit and is therefore the only one being asked about his reason to move. As migration in my analysis is restricted to economic migration, this means Heads in the Panel Study of Income Dynamics are equivalent to the decision-maker in my theoretical model. Other family members cannot be decision-makers because for them the reason to move is not know.

The Panel Study of Income Dynamics defines Head as follows: In each wave of family data "each (..) family unit has one and only one current Head."[53] Depending on the time when Head first became Head, two different types of rules are applied by the Panel Study of Income Dynamics:

1.) "Originally, if the family contained a husband-wife pair, the husband was arbitrarily designated the Head to conform with Census Bureau definitions in effect at the time the study began."[54]

2.) A new Head is selected when "last year's Head moved out of the (..) family unit, died or (…) became incapacitated; or a female Head has gotten married;"[55] or when the interviewed family is a split-off family[56] that is one or more individuals who broke away from their previous family unit and formed a new economically independent family unit.[57] In this case the following rules (…) apply: "The head of the FU [family unit, insertion by author] must be at least 16 years old and the person with the most financial responsibility for the FU. If this person is female and she has a husband in the FU, then he is designated as Head. If she has a boyfriend with whom she has been living for at least one year, then he is Head. However, if the husband or boyfriend is incapacitated and unable to fulfill the functions of Head, then the family will have a female Head."[58] Furthermore, it is possible that the female half of a pair insists on being the Head.[59]

---

[53] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 83, Paragraph 1.
[54] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 83, Paragraph 1.
[55] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 5.
[56] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 5.
[57] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3.
[58] Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 83, Paragraph 2.
[59] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 84.

4.3 Definition of migrant

4.3.1 Definition of migrants in the empirical analysis

The abstract term „migrant“ used in the theoretical model of Part A must be defined precisely for the empirical analysis. Given the limitations and definitions of the Panel Study of Income Dynamics, migrants in my empirical analysis are defined as **people who move for economic reasons together with the Head of the next wave between U.S. states between two consecutive waves of the Panel Study of Income Dynamics between 2000 and 2009.**

The first part of the definition of **economic motivated moves** results from my research interest in economically motivated migration.

The second part of the definition, that the **Head of the next wave** must be part of the moving family, is closely related to the first requirement. First, only Heads are asked about their reason to move. Second, this is only done after the move actually took place. Therefore, Head of next the period is relevant. This means, if people move without next period’s Head, it is impossible to get information about their reason to move. If the reason to move is not available, migration cannot be restricted to economic migration.

An example may illustrate the problem: Consider mother and child that move in with the child’s step-father who is head of the family after the move. In the first case, the step-father himself did not move and, therefore, won’t be asked about his reason to move. Hence, the reason why mother and child moved is not available. In the second case, the step father moved too, but not initially with mother and child. Then he might have had another reason for his move than mother and child had. Here again, it is not clear whether mother and child moved for economic reasons or not. In the third case, the step father moved together with mother and child, i.e., meets the requirement of the definition. In this case, he is asked for his reason to move and economically motivated moves can be isolated.

The third part of the definition, namely **U.S. interstate migration,** is due to methodology which can best be applied for U.S. interstate migration (see Part A, Section 3.1 for a detailed discussion).

The fourth part of the definition of **consecutive waves** results from data availability, where a new wave of data is released only every two years. Consequently, only those migrants are surveyed who have a different state of residence between two consecutive waves of data. A time interval of two years has two drawbacks. First, people who move to another place but return to their former place of residence within the two year period are not recognized as migrants. For example, when someone lives in Ohio in 2001, moves to New York in 2002, but moves back to Ohio in 2003. These moves cannot be identified as migration and, thus, reduce the number of observations. The second drawback is more critical: The migration decision of people who move more often than every two years will be misspecified. For example, if someone lives in Idaho in 2001, moves to Alabama in 2002 and finally to California in 2003, is falsely identified as a person moving from Idaho to California. This results in uncontrollable biases because the migration decision which is shown in the data (moving from Idaho to California) actually never took place. Unfortunately, there is no way to rule out these cases as the data does not include such information.

The fifth and last part of the definition is the **period of investigation which is restricted from 2000 to 2009** for three reasons: First, more recent waves have not been released yet at the time the empirical analysis of my dissertation was run. Second, the American Community Survey from which the income parameters in each year are estimated is not available before 2000. Other high density samples[60] for years earlier than 2000 do not exist. Without an estimation of the income parameters, the migration decision model of this study cannot be implemented empirically. Third, going back to earlier waves would make it more probable that the general structure of risk-attitudes in the population changes.

4.3.2 Migrant versus mover

The terms migrants and movers are often used synonymously.[61] Other authors like Rogers and Castor (1983) differentiate between (i) movers who must have been moving at least once during a given period and (ii) migrants who have a different place of residence at the end of the period.[62] In this definition every migrant is a mover, but not every mover is a migrant.

In the Panel Study of Income Dynamics only people moving between two consecutive waves of data can be identified because there is no information available about where people lived between two waves of data. This means, they are migrants and movers at the same time. Consequently, the terms migrant and mover can be used synonymously in my analysis even if the very narrow definition is applied. The reminder of my analysis employs the term migrant.

[60] High density samples do not only have a large sample size but also capture a high percentage of the population.
[61] See Farley (1998), p. 279.
[62] See Rogers, Andrei, and Luis J. Castro (1983), p. 2.

Part B – Data Set and Data Cleaning

Since a data set to perform the empirical analysis does not exist, Part B is concerned with its creation. This requires extensive data cleaning and multiple merging of different data sets which is afflicted with multiple complex problems discussed and solved in this part of my study including details on the technical implementation of the solutions.

The reminder of Part B is organized as follows: Part B, Section 1 gives an overview of the data cleaning process and presents the final data set of all migrants. Part B, Sections 2 to 4 give details on the data cleaning process. More precisely, Part B, Section 2 identifies migrants and their moving families together with their socio-economic characteristics gender, age, and education from the Panel Study of Income Dynamics. In Part B, Section 3 annual income parameters are estimated for each migrant and its particular socio-economic characteristics from the American Community Survey. Part B, Section 4 creates the final data set by merging cleaned data from Part B, Sections 2 and 3 and adding risk-free discount factors needed to model migration decisions based on planning periods longer than a year.

1 Overview of data cleaning and characteristics of the final sample

All steps of the data cleaning process are afflicted with multiple complex problems that are summarized together with their solutions in this section (Part B, Sections 1.1 to 1.3). Finally, the resulting data set on migrants and its characteristics are discussed (Part B, Section 1.4).

1.1 Overview of PSID-cleaning

This section summarizes the main problems and solutions to identify economic migrants and their socio-economic characteristics from the Panel Study of Income Dynamics (for details on this issue please refer to Part B, Section 2).

PSID-Problem 1 and solution: Identifying migrants (for details see Part B, Section 2.2)

The Panel Study of Income Dynamics consists of several Single-Year Family Files which only include the current state of residence. To identify migrants by their changing state of residence, consecutive waves of Single-Year Family Files have to be merged. Unfortunately, a variable linking one and the same family in all Single-Year Family Files does not exist. This makes merging difficult.

The solution to following migrants over time can be found in another data set, the Cross-Year Individual File. It includes a time series of annual Family Interview Numbers.[63] The latter uniquely identifies each family within each Single-Year Family Files. By merging step by step all Single-Year Family Files with the Cross-Year Individual File by the respective Annual Interview Number, changing states of residence can be identified.

PSID-Problem 2 and solution: Restricting the analysis to economic migrants (for details see Part B, Sections 2.2 and 2.3)

The reason to move is not available for all migrants, but only for those who are head of the family in the wave after the move.

This problem can only be solved first, by restricting the analysis to Heads moving for economic reasons and second, by defining the Head to be the decision-maker who maximizes his preference function based on (i) his individual income or on (ii) family income.

PSID-Problem 3 and solution: Identifying socio-economic characteristic education (for details see Part B, Section 2.3)

In contrast to gender and age, education information is usually not available in the year the move took place. The reason is that individuals are asked about their education only once when they first enter the Panel Study of Income Dynamics. In addition, a second more complex problem results from the variety of education variables which are partially scattered in different data files (Single-Year Family Files and Cross-Year Individual File), follow different coding schemes, relate to different questions and, thus, partially contradict each other. Therefore, this step of data cleaning is the most complex one.

The problem of missing education information can be solved by bringing forward data from earlier waves of data. This can be done by multiple merging of Single-Year Family Files and Cross-Year Individual File. The more complex problem of inconsistent education information is solved in five steps: First, coding schemes of all waves of data are made comparable. Second, inconsistencies of several variables relating to different questions in the same year are detected and corrected. An example is a reported bachelor's degree, while another variable indicates the person did not attend college at all. Third, an algorithm is defined that summarizes partial information of different education variables in the Single-Year Family Files in one single variable. Problems to be solved by the algorithms relate to incomparable variables on foreign and U.S. education, sorting medical, law, and

[63] Note that Annual Interview Numbers of one and the same family most certainly change every year.

honorary degrees into the Bachelor-Master system applied in all other variables, unspecified answer possibilities like “other college degrees”, and years of college attendance without a reported degree. Fourth, the summarized variable of the Single-Year Family Files is made comparable to the education variable in the Cross-Year Individual File. Fifth, six different ways to define education at the time of the move are applied in order to account for different ways to interpret the data and run a robustness check in the final analysis.

PSID-Problem 4 and solution: Identifying family ties (for details see Part B, Section 2.4.)

Family ties are needed to empirically analyze the migration decision based on family income and to detect family constellations (like getting divorced) that may interfere with pure rational behavior assumed by the decisionmodel. As the family composition often changes in the course of the move (e.g., young adults leave their parents to form their own family), a dynamic approach on family ties is needed rather than the static approach offered by the Panel Study of Income Dynamics which only contains the status before and after the move.

To solve this problem, a single variable called Move Context is created which classifies each person into a system of family ties with four different types of moving constellations: Single-moves, pair-moves, family-moves, and other moves; all of which are further categorized depending on the family ties before and after the move. The Move Context Variable takes into account all information available, such as relation by blood and adoption, step-relations, foster children, son/daughter-in-law, being a pair etc.

PSID-Problem 5 and solution: Unique identification of individuals and families (for details see Part B, Section 2.5)

The final data set should include all individuals and families who moved between 2000 and 2009. Unfortunately, the raw data set includes only annual identifiers which repeat themselves every year. For example, the family identifier is assigned based on receipts of the interview in a certain year, with the first responding family getting number 1, the second number 2 and so on.[64] Consequently, there are families numbered 1, 2 and so on in every wave. The variable that uniquely identifies each individual and family among all migrants between 2000 and 2009 is needed but missing.

A straight forward solution to this problem is to sort all migrants in ascending order by year of the move, their annual Family Interview Number and their Sequence Number and assign consecutive numbers.

[64] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3.

1.2 Overview of ACS-cleaning

This section first, presents the main problems and solution in cleaning data of the American Community Survey (for details on this issue please refer to Part B, Section 3), and second, gives a stylized overview of income parameters estimated from the American Community Survey (for detailed data please refer to data file „\American Community Survey\(12) Table of all Income Measures in all years\12_All Income Measures in all years.sav" on the data storage that comes along with this dissertation).

ACS-Problem 1 and solution: Quantification of income (for details see Part B, Section 3.2)

To estimate income parameters, the term "income" must be put in concrete terms. Given data availability and definitions of income applied by the American Community Survey, income possibilities for each U.S. state are estimated based on "total pre-tax personal income or losses from all sources for the previous year"[65] of all residence of that state. Family income is estimated based on the two main earners of the family, usually Head and Wife.

ACS-Problem 2 and solution: Reducing socio-economic groups (for details see Part B, Section 3.4)

Empirical estimation of income parameters requires a minimum sample size of at least 30 respondents for every combination on gender, age, education, U.S. state, and year (2000 to 2009).

In order to meet this minimum sample size requirement, the number of socio-economic groups by gender, age, and education must be reduced by the following three-step procedure: First, a descriptive analysis is used to gain a first insight into the data structure regarding gender, age, and education. Second, an agglomerative hierarchical cluster analysis with the Ward-Method using squared Euclidean distance measures and z-transformed variables (mean and variance of income) is employed to reduce the number of socio-economic groups. Four different cluster definitions are applied, namely a separate clustering for each year of data based on weighted und unweighted sample and a uniform clustering over all years based on weighted and unweighted samples. Unfortunately, the cluster analysis is not able to generate the minimum sample size for all socio-economic groups without clustering all people into one single group. Especially critical are sample sizes for people with higher education than a high school diploma. Since income parameters should be specific to socio-economic characteristics, the last clustering into one single group is not performed. Instead a third step is taken. That is, economic considerations are used in combination

[65] Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable "INCTOT", paragraph "Description".

with insights from the descriptive analysis to further reduce the number of groups within each of the four cluster definitions until the required sample size is reached in most groups. To understand this third step, an example might help: For individuals younger than 25, education levels Bachelor and Master are pooled. The economic consideration in this case is that the small sample size is due to the fact that the majority of people in this age group is simply too young to have a college degree. This argument is in line with insights from the descriptive analyses where mean and variance of income of bachelor and master graduates are very similar for people aged 15 to 24.

ACS-Problem 3 and solution: Estimating annual income parameters (for details see Part B, Section 3.6)

The sample sizes reached by reducing socio-economic groups are still critically small in some combinations of gender, age, education, U.S. state, and year. Note that the number of socio-economic groups cannot be further reduced without ending up with one single group only, i.e., income parameters are not specific to socio-economic characteristics anymore.

To solve this problem, annual income parameters are estimated from three different samples relating to three different time periods, namely (i) annual income reported in the year of the move, (ii) annual income reported in the three years surrounding the year of the move adjusted for inflation, and (iii) annual income reported in the three years surrounding the year of the move not adjusted for inflation. In case of weighted cluster definitions, annual income data from three years of data must be additionally re-weighted by their annual weights.

Stylized overview of income parameters estimated from the ACS

To gain an insight into the variety of income parameters for different socio-economic groups and the relation of mean and variance of income that is important for the interpretation of the results, Part C, Table 2, p. 34 gives an example of income parameters estimated from the American Community Survey. Exemplary for all income parameters estimated Table 2, p. 34 reports mean and variance and standard deviation for Head's personal income in 2009 in the U.S. state of Alabama based on pooled clustering, income data from one year weighted to be representative of the whole population.

			Annual income parameters in 2009 for U.S. state X		
Age	**Education**	**Gender**	**Expected values (μ)**	**variance of income (σ^2)**	**standard deviation of income (σ)**
15-24 years old	**Less than high school**	**Male**	2,478	27,521,155	5,246
		Female	1,483	10,653,051	3,264
	High school/ Associate	**Male**	10,351	136,942,841	11,702
		Female	7,736	127,591,397	11,296
	Bachelor/Master and higher		18,341	258,851,924	16,089
25-34 years old	**Less than high school**		11,754	171,587,648	13,099
	High school/ Associate	**Male**	28,492	577,376,817	24,029
		Female	17,610	349,411,146	18,693
	Bachelor's degree	**Male**	45,599	936,374,208	30,600
		Female	29,259	427,417,995	20,674
	Master/Profess-ional or higher	**Male**	68,330	2,773,030,630	52,660
		Female	40,506	794,652,269	28,190
35-45 years old	**Less than high school**		15,623	441,923,603	21,022
	High school/ Associate	**Male**	37,031	979,223,471	31,293
		Female	20,939	459,727,014	21,441
	Bachelor's degree	**Male**	74,964	3,750,558,562	61,242
		Female	38,058	1,075,016,097	32,787
	Master/Profess-ional or higher	**Male**	101,799	7,619,428,610	87,289
		Female	52,640	2,004,418,530	44,771
46-65 years old	**Less than high school**		16,006	510,998,654	22,605
	High school/ Associate	**Male**	39,575	1,430,753,974	37,825
		Female	21,435	508,480,186	22,550
	Bachelor's degree	**Male**	83,326	5,648,434,319	75,156
		Female	39,392	1,687,306,787	41,077
	Master/Profess-ional or higher	**Male**	115,743	9,409,994,681	97,005
		Female	56,379	2,130,459,658	46,157
Retired 66 and older	**max. Associate**	**Male**	28,430	1,011,519,880	31,804
		Female	16,366	358,550,960	18,935
	Bachelor/Master and higher	**Male**	72,415	4,532,753,627	67,326
		Female	35,611	1,439,526,766	37,941

Table 2: Example of estimated income parameters from the American Community Survey exemplary for Head's individual income in 2009 in the U.S. state of Alabama based on pooled clustering, one year's data weighted to be representative of the population of that state.

In total there are altogether 24,480 tables of income parameters like Table 2, p. 34 that are estimated from the American Community Survey resulting from 2 risk-measures (variance and semi-variance) multiplied by 2 levels of income (Head's individual and family income) multiplied by 51 destination states[66] multiplied by 4 types of clustering socio-economic groups (see ACS-Problem 2 and solution) multiplied by 3 different data sets from which income parameters are estimated (see ACS-Problem 3 and solution) multiplied by 10 years (2000 to 2009). Since reporting all numerical results on over 24,480 pages full of tables does not add to the understanding of my study, I kindly ask the interested reader to refer to file „\American Community Survey\(12) Table of all Income Measures in all years\12_All Income Measures in all years.sav" on the data storage that comes along with this study.

1.3 Overview of estimating risk-free discount factors

This section is concerned with risk-free discount rates that are needed to estimate income parameters not only for a planning period of one year but also for the period until reaching full retirement age and until reaching life expectancy (for details on this issue please refer to Part B, Section 4.3).

Problem: Gaining four input parameters needed for risk-free discount factors (for details see Part B, Section 4.3)

In order to estimate the present value of lifetime income (working life income, respectively), migrant-specific risk-free discount factors must be derived. Therefore, four input parameters are needed as follows: (ii) the time until reaching full retirement age, 2) the expected remaining lifetime, and 3) the risk-free term structure at 4) the time of the move.

Solution for 1) time until reaching full retirement age (for details see Part B, Section 4.3.1)

The remaining years until reaching full retirement for each cohort are determined by U.S. legislation. The respective full retirement ages for each cohort are taken from the U.S. Social Security Administration.[67]

Solution for 2) expected remaining lifetime (for details see Part B, Section 4.3.1)

The expected remaining lifetime of U.S. citizens for each year is documented in the United States National Vital Statistics Reports issued by the National Center for Health Statistics.[68] For years 2000

[66] The Panel Study of Income Dynamics from which migrants are identified differentiates between 50 U.S. states and Washington D.C.

[67] For legal retirement ages depending on the year born, see U.S. Social Security Administration (2011).

[68] For life expectancy in 2000 see Arias, Elizabeth (2002), pp. 9-12. For life expectancy in 2001 see Arias, Elizabeth (2004a), pp. 9-12. For life expectancy in 2002 see Arias, Elizabeth (2004a), pp. 9-12. For life expectancy in 2003 see Arias, Elizabeth (2006), pp. 10-13. For life expectancy in 2004 see Arias, Elizabeth (2007), pp. 10-13. For life expectancy in 2005 see Arias, Elizabeth, Brian L. Rostron, and Betzaida Tejada-Vera

to 2007 life expectancy is given in detail for all ages from 0 to 100 by gender. Unfortunately, data for years 2008 to 2009 is aggregated to age groups. To solve this inconsistency in the data structure, data for 2008/09 for each age is approximated by linear interpolation.

Solution for 3) risk-free term structure (for details see Part B, Section 4.3.2)

Discount factors can be determined from the term structure of interest rates. Unfortunately, such a term structure is not available for the U.S. but must be derived. For maturities up to 30 years, results are published in the paper of Gurkaynak, Sack, and Wright (2006). For maturities over 30 years, I derive the term structure on an annual basis using the same methodology as Gurkaynak, Sack, and Wright (2006), i.e., the extension of the Nelson-Siegel (1987) approach proposed by Svensson (1994).

Solution for 4) time of the move (for details see Part B, Section 4.3.2)

The term structure can be determined on a daily basis and hence might change from day to day. In contrast, the exact date of the move is not available for most migrants but only the year of the move. Since the move could have been taking place during all days of the year, it is not clear which day's term structure is relevant for the move. This means, the frequency of migration data (annual) and data on term structure (daily) is not congruent. To solve the problem, annual term structures are needed that are congruent to the annual data on moves. They are derived by simply taking the average of daily term structures of the respective year. Note that this approach is in line with the way income is reported in the American Community Survey: annual income includes income of the last 12 month because interviews are run throughout the year.[69]

1.4 Key characteristics of the final sample of migrants

The final sample includes 321 migration decisions that can be empirically analyzed. The whole sample can be examined with respect to Head's personal income, but only 315 of them based on family income. The reduced sample size of family decisions compared to individual decisions is due to missing socio-economic group definitions for the second main earner of some families, usually due to missing education information. The 315 families surveyed account for a total of 820 family members.

To get an impression of the key characteristics of the sample, the reminder of this section briefly presents frequency and distribution of the key characteristics of the sample with respect to (i) the

---

(2010), pp. 10-13. For life expectancy in 2006 see Arias, Elizabeth (2010), pp. 9-12. For life expectancy in 2007 see Arias, Elizabeth (2011), pp. 10-13. For life expectancy in 2008 see Miniño, Arialdi M., Sherry L. Murphy, Jiaquan Xu, and Kenneth D. Kochanek (2011), p. 74, Table 6.

[69] See Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable „INCTOT", paragraph on "Description".

year of the move, (ii) Head's and Wife's socio-economic characteristics gender, age, and education, (iii) family size, (iv) families ties and (v) the special case of family ties: divorce/separation.

<u>Number of observation by year between 2000 and 2009</u>

Although unlikely, it could be that risk-attitudes systematically differ over years, for example, rising risk-aversion over years. To avoid such a bias, it is important to have a about the same number of observations over the investigation period from 2000 to 2009. Figure 2, p. 37 illustrates the frequency distribution of the 321 moves by years. The relative low number of moves in 2009 is due to the fact that the last wave included in this analysis is the one of 2009. As interviews are run throughout the year, family data of 2009 does not include all moves of 2009, but only those which took place before the interview. Like in all other waves included in the survey, a considerable number of moves are reported only in the next wave of data.

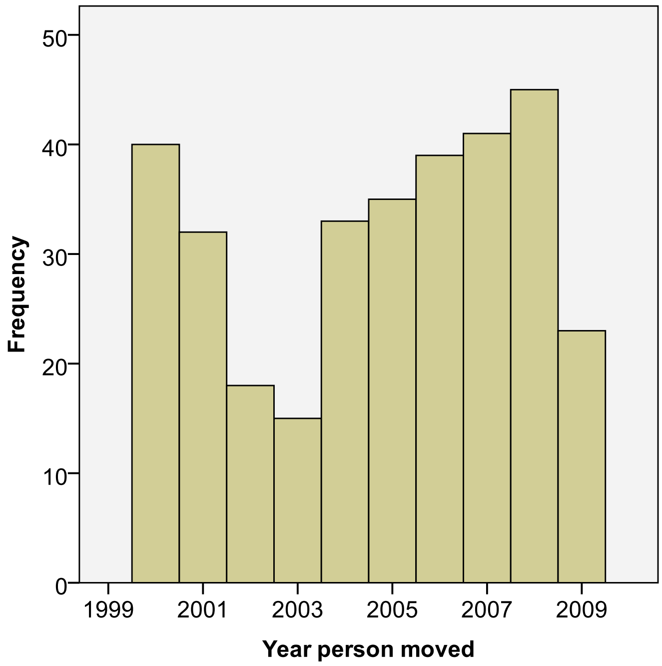


Figure 2: Histogram of years at which migration took place in the personal migration sample.
Source: Own illustration based on own calculations.

<u>Socio-economic characteristics gender, age, and education</u>

To be able to estimate the effect of socio-economic characteristics gender, age, and education on risk-attitudes, it is important to have considerable variation in these variables. In both samples the vast majority of Heads is male (82%, see Table 3, p. 38) with men being on average 5 years older (average age of male Heads 35) and slightly less educated than their female counterparts (see Table

4, p. 38). Among individuals being the second main earner, it is again women having a higher education level (see Table 4, p. 38).

	Personal migration sample	Family migration sample
Total number of Heads	321	315
Number of male Heads (percentage of male Heads)	264 (82%)	259 (82%)
Number of female Heads (percentage of male Heads)	57 (18%)	56 (18%)

Table 3: Sample size by gender in the personal and family migration samples.
Source: Own illustration based on own calculations.

	Mean age		Mean education	
	Male	**Female**	**Male**	**Female**
Personal migration sample (n=321)				
Heads	35	30	4.13	4.39
Family migration sample (n=315)				
Heads (ranked highest)	35	30	4.14	4.36
Second main earner (ranked second highest)	16	35	3.44	4.11

Table 4: Mean age and mean education of Heads and Wives in the personal and family migration samples.
Source: Own illustration based on own calculations. Education definition one is applied (see Part B, Section 2.3.3.6.4 for a detailed definition). All other education definitions look alike.
Education level 3 denotes "High school graduate/GED", 4 denotes "Associate's degree", 5 denotes "Bachelor's degree".

The frequency distribution of ages of all Heads in the samples is illustrated in Figure 3, p. 39 exemplary for the personal migration sample (family migration sample looks alike). It shows that there is a slight tendency of lower number of observations for older migrants. Ages rank from 19 to 78, with 99% of all 321 Heads being between 20 and 60 years old.

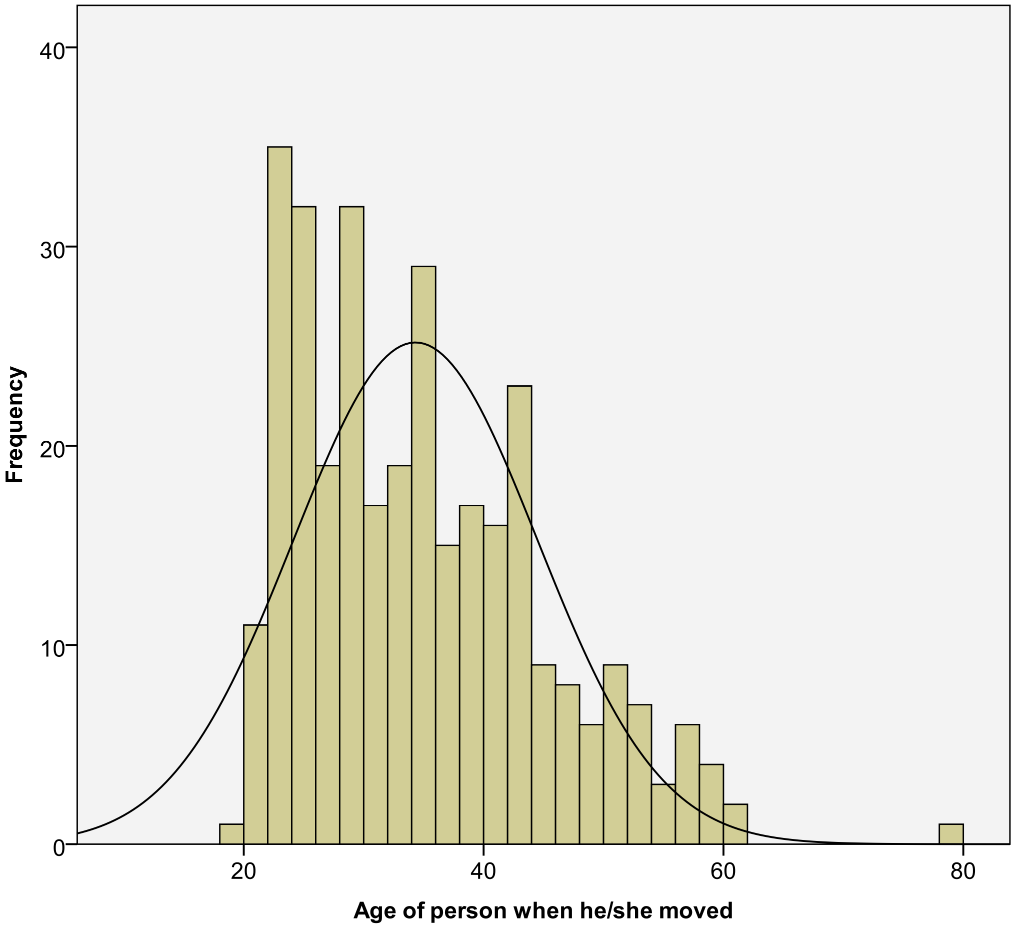


Figure 3: Frequency distribution of Heads' age against the normal distribution exemplary for the personal migration sample.
Source: Own illustration based on own calculations. Sample size 321, mean age 34.32, standard deviation 10.17, modus 29, median 33.

Figure 4, p. 40 provides information about the education levels of Heads exemplary for the personal migration sample (the family migration sample looks alike). Overall, 92% of Heads have a high school diploma or a higher degree up to a master's degree irrespective of gender. Vice versa, very low (less than high school) and very high levels of education (Doctorate, Ph.D.) are rare among Heads in the samples. The low frequency of Heads with a Doctorate or Ph.D. (7 Heads) does not result from the age distribution in the sample, but is due to occurrence of these in the population. Note that more than 59% of the sample are older than 29 – an age at which a Ph.D. is possible.

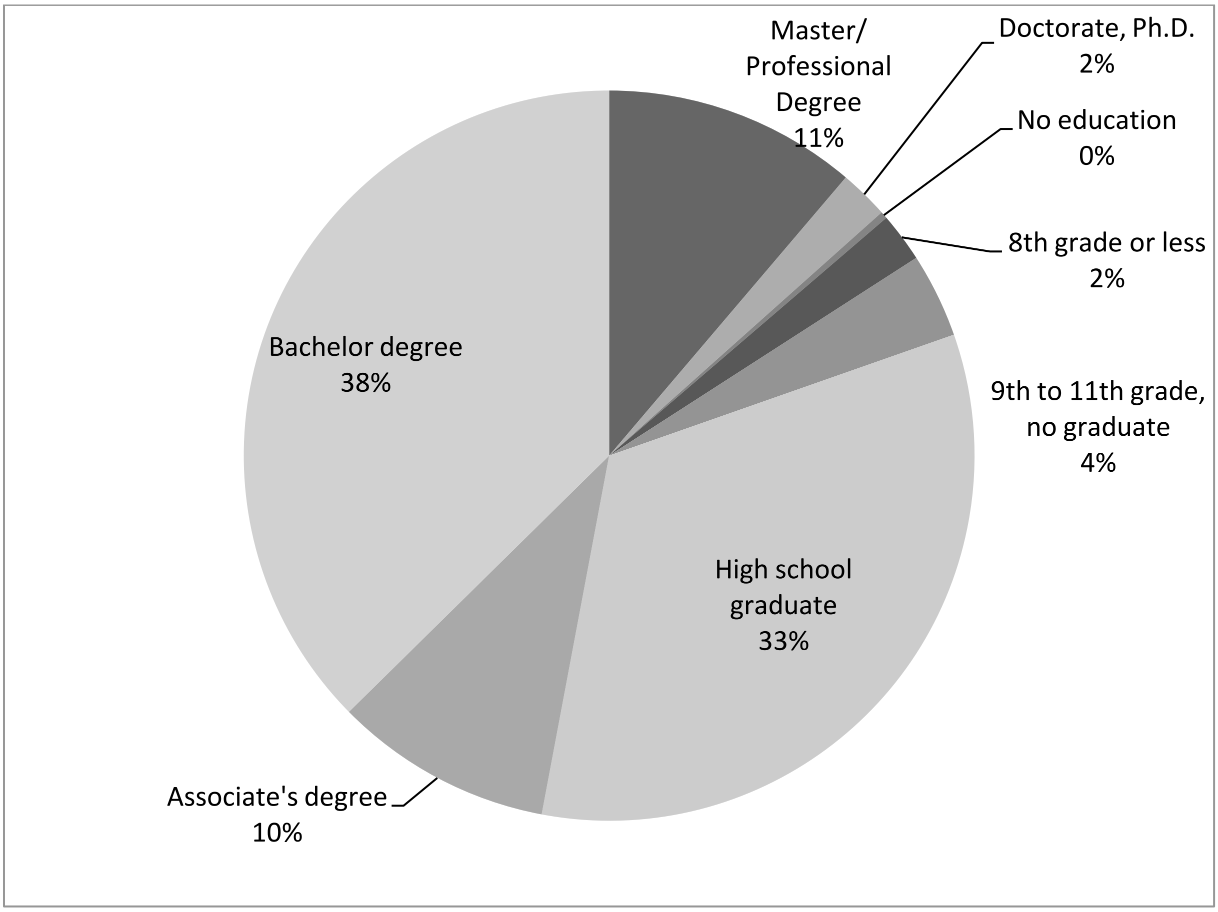


Figure 4: Education levels of Heads exemplary for the personal migration sample.
Source: Own illustration based on own calculations. Education definition one is applied (see Part B, Section 2.3.3.6.4 for a detailed definition). All other education definitions and the family migration sample look alike.

Obviously, there is enough variation in the sample concerning gender, age, and education to gain reliable estimations of their respective effect on risk-attitude.

Number of people moving together as a family

The number of people moving together as a family might relate to the risk-attitude of the decision-maker. Hence, a certain variation on the number of people moving together is again required to estimate this effect. The average number of people moving together as family is 2.39. The frequency distribution of the family sizes at the time of the move, before and after the move ranking from one-person families to 10-person families is provided in Table 5, p. 41.

Number of family members …	… actually moving together		… before the move		… after the move	
	Frequency	**Percentage**	**Frequency**	**Percentage**	**Frequency**	**Percentage**
1	133	42.2	72	22.9	108	34.3
2	55	17.5	92	29.2	72	22.9
3	44	14.0	52	16.5	52	16.5
4	52	16.5	61	19.4	49	15.6
5	20	6.3	23	7.3	21	6.7
6	8	2.5	10	3.2	10	3.2
7	3	1.0	4	1.3	3	1.0
10	0	.0	1	.3	0	.0
Total	**315**	**100.0**	**315**	**100.0**	**315**	**100.0**

Table 5: Average family size at the time of the move, before and after the move.
Source: Own illustration based on own calculations.

Table 5, p. 41 reveals that there are actually not only single movers but also families with up to 10 people that live together. Furthermore, the frequency distribution of the number of people living together before/after the move compared to those actually moving together already suggests that the explanatory power of the corresponding variables on risk-attitudes might be different.

<u>Family ties</u>

The effect of family ties, namely single-moves, pair-moves, and family-moves can only be investigated if these constellations are available in the data set. Therefore, Table 6, p. 42 provides an overview of family ties of the 315 Heads in the family migration sample. The sample is relatively balanced between one-person moves (single moves accounting for 43%) and several-person moves (pair- moves and family-moves accounting for 57%). Within each category of family ties the typical constellation is always the most common: In 76% of all single-moves Head lives alone before and after the move; 99% of all pair-moves are typical partner moves, where two people move together as a pair (48 families) sometimes having newborns (9 families). The same picture can be drawn for the family moves where 85% move in an unchanged family constellation (80 families) sometimes getting babies (24 families).

	Number of families
Single-move	**134 (43% of 315 families)**
Moves alone and lives alone now (not leaving partner)	102
Head leaves partner	8
Moves in with parent/partner/child/other relatives (not leaving partner)	14
Moves in with unknown person with no relation	10
Pair-move	**58 (18% of 315 families)**
Two people move together being a pair before or/and after the move	48
Pair with newborn(s)	9
Moves in with parent/partner/child/other relatives (not leaving partner)	1
Family-move	**123 (39% of 315 families)**
Whole family moves unchanged	80
Whole family moves plus newborn(s)	24
Part of family moves (not leaving partner)	15
Part of family moves, leaving partner	4

Table 6: Family ties of Heads in the family migration sample.
Source: Own illustration based on own calculations.

<u>**Separation from old partner**</u>

A family constellation that might interfere with the pure economic reason to move is the separation of Head from the old partner. Table 6, p. 42 shows that among all 315 family moves only 12 Heads (4%) separated from their old partner of which 8 Heads lived alone after the separation, and 4 Heads are joined by a part of their old family. The low percentage of separated Heads already suggests that the investigation of the relation of separation and risk-attitude might be problematic due to few observations in the sample.

2 Data cleaning of the Panel Study of Income Dynamics

This section is concerned with data cleaning of the first data source, the Panel Study of Income Dynamics. The data cleaning is performed via SPSS programming. The detailed coding algorithm can be found in the SPSS Syntax Files and its commands. The names of the corresponding SPSS-Syntax files and resulting SPSS data files are given in footnotes at the beginning of each section.

2.1 Overview: problems, solutions, and algorithm

The objective of the first step of data analysis is to identify single migrants as well as families that move for economic reasons. Furthermore socio-economic characteristics gender, age and education of all family members must be identified. This is associated with several problems that are outlined below.

General Problem 1: Identifying movers

Movers can be identified by comparing their state of residence for two consecutive waves of data. Family files only contain the current state of residence. Cross-Year Individual File do not contain any information concerning residence. This raises the question of how to find out about a person's former state of residence.

General Problem 2: Economic movers

The decision model derived in Part A only holds for economic migrants. Therefore, the sample of migrants must be restricted to economic migrants.

General Problem 3: Identifying socio-economic characteristics

In order to estimate income parameters specific to the socio-economic characteristics of the migrant, information on gender, age, an education of each migrant are needed. Variables regarding gender and date of birth of every person are available in the Cross-Year Individual File. Information on education is available in both types of data sets, namely the Single-Year Family Files and the Cross-Year Individual File, but their information content differs a lot. While the Cross-Year Individual File gives the number of school years completed up to 17 years for every person, the Single-Year Family Files contain education information only on Heads and Wives, but in more detail including the highest academic degree achieved. This raises the question of how to combine education variables of different information content and in different types of data files.

Additionally, the problem of education variables in the Single-Year Family Files is that they belong to the category of so called "background information" variables.[70] Questions belonging to this category are only asked once for every person - in the year when Head/Wife first became Head/Wife. This results in missing data for Head's background information, if Head is not new in this family and missing background information for Wife if Wife is not new in this family. Consequently, variables about Head's/Wife's education are only valid, if the family has a new Head/Wife. As a consequence, education variables must be brought forward from earlier waves of family data. Because of enormous changes in the data structure over the years this raises multiple problems.

General Problem 4: Identifying family ties

Family ties are important for two reasons: First, to model the migration decision based on family income, it is necessary to know who is part of the family. Second, personal relationships such as getting divorced may interfere with pure economic reasons to move. Unfortunately, figuring out whose is part of the family at the time of the move is a tricky task as family compositions usually change during the migration process (e.g., parents get divorced, children form their own economically independent family, grandparents move in with their children and so on). The Panel Study of Income Dynamics does not contain detailed information about family ties before and after the move. This leaves the question of how to identify family ties between two consecutive waves of data.

General Problem 5: Unique identification of individuals and families

The final data set consists of all individuals and families, respectively who moved between 2000 and 2009. Unfortunately, the raw data set includes only annual identifiers which repeat themselves every year. For example, the family identifier is assigned based on receipts of the interview in a certain year, with the first responding family being numbered 1, the second 2 and so on.[71] Consequently, there is a family numbered 1, 2 and so on in every wave, but this family most certainly is not the same family as the one with the same number in the previous or next year. Thus, a unique identifier for the final data set is missing.

General solution to these problems

The principal idea is to get a single file with one record for each migrant that includes all individual and family information needed for my analysis over several years of data. This can be achieved by

[70] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 85, Paragraph 1.
[71] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3.

multiple merging of Single-Year Family Files and the Cross-Year Individual File wave by wave. To solve the Problems 1 to 4 outlined above the following steps have to be taken:

1.) By merging of consecutive waves of Single-Year Family Files and the Cross-Year Individual File economic migrants are identified (see Part B, Section 2.2).
2.) Socio-economic characteristics on gender, age and education are identified in Part B, Section 2.3. The definition of education at the time of the move is afflicted with multiple complex problems for which I refer to Part B, Section 2.3.3.1.
3.) Family ties of each migrant are identified in Part B, Section 2.4.
4.) Finally, variables that uniquely identify each migrant and migrant family in the final data set which collects migrant from all waves of data are added (see Part B, Section 2.5).

2.2 Identifying economic migrants[72]

2.2.1 Problems and solution

General Problems 1 and 2 of Part B, Section 2.1 consist of the following two subproblems.

Problem 1: Identifying migrants

The total migrant stock surveyed in my study is composed of five waves of migrants, namely those who moved between 1999/2001, 2001/2003, 2003/2005, 2005/2007, and 2007/2009. Migrants of each wave can be identified by their changing state of residence from one wave to the other. Unfortunately, the Single-Year Family Files of each wave only contain the current state of residence which is not enough to identify migrants.

Problem 2: Reason to move only available for Heads

Once all moving individuals are identified, only those moving for economic reasons must be selected in order to fulfill the assumption of the migration decision model estimated in this study. This is problematic as only Heads are asked about their reason to move. All other family members are not asked about their reasons.

[72] The SPSS Syntax file corresponding to this section is named "Identify_Mover_07_09 (1-6).sps" and can be found separate for each migrant wave in folder „\Mover\(1) Mover between Waves".

Solution to Problem 1: Identifying migrants

By merging family data of consecutive waves migrants can easily be identified by their changing state of residence from wave to wave.[73]

Solution to Problem 2: Reason to move only available for Heads

Because the reason to move is only available for Heads, he is defined to be the decision-maker in this analysis, who maximizes his preference function based on (i) his own income or (ii) on the family income. If people move together as a family, it can be assumed that the reason given by the head of the family applies to the family as a whole. In contrast, if someone who is not Head moves with people who become only lower level family members (not Head) after the move, there is no variable indicating his reason to move since Head is the only one who is asked about it. The individual must be excluded from this study as the reason to move is unknown on the individual as well as the family level.

If someone moves in with a new family who's Head has accidentally moved, too, but not together with the individual in question, Head's reason to move may be different to the reason of the individual in question. In this case, it cannot be differentiated between economic and non-economic migrants. Therefore, my analysis is restricted to migrants that move together with Head.

2.2.2 SPSS-implementation[74]

In order to identify economic migrants data of the Cross-Year Individual File is complemented by family variables of the waves before and after the move. The merging variable is the corresponding annual Family Interview Number. The family variables added are:

- state of residence before and after the move (STD_StateOrigin, STD_StateDestination)
- reason to move in the year after the move.[75]

[73] Steps 1 to 3 in the SPSS Syntax file „\Mover\(1) Mover between Waves\2007_2009\ Identify_Mover_07_09 (1-6).sps".

[74] The variables mentioned here are added to the SPSS-data file "\Panel Study of Income Dynamics\ 30_Collected EDU of ALL 89-09.sav". The SPSS Syntax corresponding to this section are named "Identify_Mover_07_09 (1-6).sps" and can be found separate for each migrant wave under „\Mover\(1) Mover between Waves".

[75] See steps 1 and 2 in the SPSS Syntax file (exemplary for migrant wave 2007/2009): „\Mover\(1) Mover between Waves\2007_2009\ Identify_Mover_07_09 (1-6).sps". Besides the state of residence and the reason to move, the number of people living in the family before and after the move is also added for two reasons: First, the number of people in the family may influence the risk-attitude estimated in the family context. Second, it makes the classification of family ties more easily as will be discussed in Part B, Section 2.4.

Next, only those who moved between U.S. states and together with the later Head are selected.[76] Finally, only those moving for economic reasons are selected.[77] The resulting data sets includes all U.S. interstate migrants moving for economic reasons with all their personal data (from the Cross-Year Individual File) and all education data from the Cross-Year Individual File and the Single-Year Family Files.

2.3 Identifying socio-economic characteristics

This section refers to general Problem 3 outlined in Part B, Section 2.1. In order to estimate income parameters specific to the socio-economic characteristics of the migrant, information on gender, age, an education of each migrant is needed. Unfortunately, this is afflicted with several problems that are discussed and solved in this section.

2.3.1 Gender

Gender of each person ever in the Panel Study of Income Dynamics is available in the Cross-Year Individual File. It is assumed that a person's gender does not change during life time. Hence, gender reported once when the person first entered the Panel Study of Income Dynamics also holds for the time of the move. This seems to be in line with the Panel Study of Income Dynamics assumptions, as there is only one variable on gender in the whole data set. If the variable would account for people changing their gender, there would be an annual gender variable.

2.3.2 Age

As the Single-Year Family Files include the month and year of the move (variables STD_MonthMoved, STD_YrMoved) and the Cross-Year Individual File includes the month and year born (variables STD_YrBorn, STD_MonthBorn), age at the time of the move can easily be derived by:

STD_AgeMoved = STD_YrMoved - STD_YrBorn.

If the person was born in the same month he moves or even later, STD_AgeMoved is reduced by one year. The intuition underlying this algorithm is that the decision about moving was taken before the

[76] See steps 3 to 5 in the SPSS Syntax file (exemplary for migrant wave 2007/2009): „\Mover\(1) Mover between Waves\2007_2009\ Identify_Mover_07_09 (1-6).sps". The resulting data file is named „\Mover\(1) Mover between Waves\2007_2009\4_all_mover_with_Head_05_07.sav". Auxiliary SPSS data files with temporary results are named "0_CollectedEDU_sorted.sav", "1_CollectedEDU_09StateReason.sav", "2_all mover_07_09.sav", and "3_moving Heads_07_09.sav".

[77] See step 6 in the SPSS Syntax file (exemplary for migrant wave 2007/2009) „\Mover\(1) Mover between Waves\2007_2009\ Identify_Mover_07_09 (1-6).sps". The resulting data file is named „\Mover\(1) Mover between Waves\2007_2009\ 5_all econ_mover_07_09.sav".

move actually takes place. An example may clarify the idea: Somebody born in August 1950 and moving in August 1990 is assumed to be 39 at the time of the move. It seems reasonable to assume that the decision about moving was taken before August 1990. This procedure results in negative ages for those who are born only after the move. An age of minus one implies that the person was born in the year after the move. For those born in the month the move takes place, the age is corrected to zero, in order to be able to differentiate between those only born after and those born “during” the move.

2.3.3 Education

2.3.3.1 Overview of problems and solution

Problems

Usually education at the time of the move can be figured out by simply looking at the education reported for the respective year. In case of the Panel Study of Income Dynamics this is afflicted with multiple problems which will be discussed in detail in the following sections: First, the Single-Year Family Files contain various variables on education. Second, the Cross-Year Individual File includes one more variable on education for each year. Third, education variables on both sources are not comparable or may even contradict each other as they relate to different questions, thus including different information. Fourth, it could happen that no education information is available in a certain year despite numerous variables on education exist. This is due to the structure of the Panel Study of Income Dynamics where education belongs to background information that is only asked once when the person first enters the Panel Study of Income Dynamics (see Part A, Section 3.4.1.2.4).

Solution

In order to define education at the time of the move, several steps must be taken, which are discussed in detail in the following subsections:

1.) The 10 education variables in the Single-Year Family Files have different coding schemes from wave to wave. Therefore, education variables must be made comparable over years by creating a standardized variable with a uniform coding scheme over all waves (see Part B, Section 2.3.3.2).

2.) Because education variables are now standardized, the information of the 10 education variables in the Single-Year Family Files can be summarized in a one single variable with the same algorithm for each year of data (see Part B, Section 2.3.3.3).

3.) All education variables of all years and all sources (Single-Year Family Files and Cross-Year Individual File) are collected in a one single file in order to simplify further steps of programming (see Part B, Section 2.3.3.4).

4.) The summarized education variable of the Single-Year Family File must be made comparable to the education variable of the Cross-Year Individual File which includes different education levels (see 2.3.3.5).

5.) Even when education variables of all sources (Single-Year Family Files and Cross-Year Individual File) are comparable across all years of data, they can still contradict each other. Therefore, an algorithm must be found to finally define education at the time of the move (see Part B, Section 2.3.3.6).

2.3.3.2 Standardize education variables over all waves of data[78]

The standardization procedure is exactly the same for variables on Head's as well as Wife's education because the questions asked are exactly the same for Heads and Wives within each wave. Therefore, in this section the standardization procedure is described only for Head's education variables.

2.3.3.2.1 Problems and solution

Different coding schemes across all waves of Single-Year Family Files must be made comparable. Two types of problems occur: Different possible answers to one and the same question (see Problem 1 below) and different coding of one and the same answer (see Problem 2 below).

Problem 1: Different possible answers

The first type of problem refers to different answer possibilities. Here is an example of different codes for the same question "Did you graduate from high school, get a GED or neither?"

[78] Corresponding SPSS-Syntax files of, for example, wave 2009 can be found under „\Panel Study of Income Dynamics\fam2009er US Haushaltspanel (PSID)\fam2009er_data Standardize (1-4).sps" concerning Head's education and "„\Panel Study of Income Dynamics\fam2009er US Haushaltspanel (PSID)\fam2009er_data Standardize (5).sps" concerning Wife's education. Files and folders for all other waves are named analogously.

- **Single-Year Family File 2001**[79]

    0 Inap.: Head is not new; educated outside US only; no education

    1 Graduated from high school

    2 Got a GED

    3 Neither

    4 Wild Code[80] (should be NA)

    8 DK

    9 NA, RF

- **Single-Year Family File 2007**[81]

    0 Inap.: educated outside US only; no education

    1 Graduated from high school

    2 Got a GED

    3 Neither

    4 College level only

    9 NA, DK

Where Inap. denotes inapplicable[82] , GED denotes General Education Development Test which certifies high school level academic skills, DK denotes don't know , NA denotes not ascertained, RF denotes refused[83].

In both years only answers coded 1 to 3 really answer the question of whether someone graduated from high school, got a GED or neither. All other answers (codes 0, 4, 8 and 9) refer to different categories that do not answer the question asked. They are therefore irrelevant to this study.

Problem 2: Different coding of same answers

The second problem refers to questions where the same possible answers are coded differently in different waves. An example is the question "What was the highest degree or certificate you (HEAD) earned outside the U.S.?" The following coding schemes can be found in the raw data.

---

[79] See Institute for Social Research, Survey Research Center, University of Michigan (2008f), p. 382.
[80] In survey research, wild codes are codes that are not authorized for a particular question. For instance, if a question that records the sex of the respondent has documented codes of "1" for female and "2" for male and "9" for "missing data," a code of "3" would be a "wild" code, sometimes called an "undocumented code". International Labour Organization (2004), p.55.
[81] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 1355.
[82] Inapplicable is given when "the value of the attribute cannot be determined for the concerning record, simply because it does not exist" (Matthé and Tré (2010), p. 141). In this context an example is when a person is only educated outside the U.S. and is asked for his high school diploma.
[83]See Andreski, McGonagle, Schoeni (2007), p. 2.

- **Single-Year Family File 1999**[84]

  0 Inap.: Head is not new; educated in the U.S. only or no education; no certificate or degree

  1 Sixth grade diploma/certificate

  2 Grammar/elementary/grade school diploma/certificate; 9th grade diploma

  3 High school diploma

  4 Associate

  5 Bachelor

  6 Master

  7 PhD; MD; DD

  9 NA ;DK; refused

- **Single-Year Family File 2007**[85]

  0 Inap.: none or did not complete grammar/ elementary/primary school; educated in the U.S. only

  1 Completed grammar/elementary/ primary school but no secondary or high school

  2 Started secondary or high school but did not finish

  3 Secondary or high school diploma

  4 Associate's degree/teaching license

  5 Bachelor of Arts/Science/ Letters; BA; BS

  6 Master of Arts/Science; MA; MS; MBA

  7 Doctorate; PhD

  9 NA; DK

Where Inap. denotes inapplicable, PhD denotes research doctorate or the Doctor of Philosophy,[86] MD denotes Medical Doctor,[87] DD denotes Doctor of Divinity,[88] NA denotes not ascertained and DK denotes don't know.[89]

Here a possible answer in both waves is "I completed grammar school" which is coded 2 in the 1999 wave and coded 1 in the 2007 wave.

---

[84] See Institute for Social Research, Survey Research Center, University of Michigan (2008e), p. 838.
[85] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 1364.
[86] See U.S. Department of Education, International Affairs Office (2008a), p. 1 and 3.
[87] See U.S. Department of Education, International Affairs Office (2008b).
[88] See U.S. Department of Education, (1995), p.3.
[89] See Andreski, Patricia, Katherine McGonagle, and Robert Schoeni (2007), p. 3.

Solution

The two problems outlined above can both be solved by creating additional "standardized variables" that have a uniform coding scheme across all waves of data. This means, they have the same answer possibilities that only refer to relevant answers (Problem 1) and these answers are coded with the same numbers across all waves (Problem 2). An example may clarify the idea.

A good example for the creation of standardized variables is the variable labeled "WTR GRADUATED HS-HD" which belongs to the question "Did you [Head] graduate from high school, get a GED or neither?" and has the following coding schemes in Single-Year Family Files 2001 and 2007.

- **Family File 2001**[90]

  0 Inap.: Head is not new; educated outside US only; no education

  1 Graduated from high school

  2 Got a GED

  3 Neither

  4 Wild Code (should be NA)

  8 DK

  9 NA, RF

- **Family File 2007**[91]

  0 Inap.: educated outside US only; no education

  1 Graduated from high school

  2 Got a GED

  3 Neither

  4 College level only

  9 NA, DK

Obviously answer possibilities vary (Problem 1) as well as their coding (Problem 2). The standardized variable to this question is defined as listed in Table 7, p. 53.

[90] See Institute for Social Research, Survey Research Center, University of Michigan (2008f), p. 382.
[91] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 1355.

Question: “Did you [HEAD] graduate from high school, get a GED, or neither?” Variable Label : "Standardized: WTR GRADUATED HS-HD"	
Parameter Values	**Value Labels**
1	"Graduated from high school"
2	"Get GED"
3	"Neither"
9999	"Missing"

Table 7: Coding scheme of standardized variables “Whether graduated from high school.”
Source: Own table.

The uniform coding scheme only includes answers which give a straight forward answer to the question. Applied to the original variables given above, relevant answers as defined in the uniform coding scheme are coded 1 to 3 in both waves and are thus transferred to the standardized variable without change. All other codes are summarized by the default value 9999 and are labeled “missing” for different reasons. First, answers like “don’t’ know” (DK) and codes like “not ascertained” (NA), “refused” (RF) or a Wild Codes do not give an answer to the question and can thus not be distinguished in their meaning regarding the question. Second, inapplicable (Inap.) answers do not contain a straight forward answer to the question (therefore coded inapplicable). Furthermore, these answers could have different meanings which can also be obtained from other variables. For example, whether Head is new or not and whether Head was educated outside the U.S. only or did not receive any education is asked in separate questions. Therefore, inapplicable answers do not add any value regarding the question of interest and no information is lost by recoding them to default values. The same argument holds for the answer “college level only” coded with 4 in the 2007 wave. This answer does not answer the question about high school. All these irrelevant answers which do not contain any information regarding the question asked are summarized and classified as “Missing” in the SPSS programming. The same logic is used for all the other standardized variables.

2.3.3.2.2 List of variable that must be standardized

The problem of different coding schemes concerns almost all variables of background information. Relevant for my study are only variables of Head’s and Wife’s educational attainment, which are listed below with their respective questions and official labels in the SPSS Single-Year Family Files. These variables will be standardized across all waves. Please remember that questions about Head’s and Wife’s education are the same.

Questions asked	SPSS Variable Label
Whether Family Unit has a new Head in this year.	CKPT: WTR NEW HEAD IN FU
Where did you [HEAD] receive your education?	WTR HEAD EDUCATED IN US
Did you [HEAD] graduate from high school, get a GED, or neither?	WTR GRADUATED HS-HD
How many grades of school did you [HEAD] finish?	GRADE OF SCHOOL FINISHED-HD
Did you [HEAD] attend college?	WTR ATTENDED COLLEGE-HD
What is the highest year of college you [HEAD] have completed?	HGHST YR COLL COMPLETED-HD
Did you [HEAD] receive a college degree?	WTR RECD COLLEGE DEGREE-HD
What is the highest college degree you [HEAD] have received?	HGHST COLLEGE DEGREE RECD-HD
How many years of school did you [HEAD] complete outside of the U.S.?	YRS FOREIGN EDUCATION-HD
What was the highest degree or certificate you [HEAD] earned outside the U.S.?	FOREIGN DEGREES-HEAD
Whether Family Unit has a new Wife in this year.	CKPT: WTR NEW WIFE IN FU
Where did your [wife/"WIFE"] receive her education?	WTR WIFE EDUCATED IN US
Did she graduate from high school, get a GED, or neither?	WTR GRADUATED HS-WF
How many grades of school did she finish?	GRADE OF SCHOOL FINISHED-WF
Did she attend college?	WTR ATTENDED COLLEGE-WF
What is the highest year of college she has completed?	HGHST YR COLL COMPLETED-WF
Did she receive a college degree?	WTR RECD COLLEGE DEGREE-WF
What is the highest college degree she has received?	HGHST COLLEGE DEGREE RECD-WF
How many years of school did she complete outside of the U.S.?	YRS FOREIGN EDUCATION-WF
What was the highest degree or certificate she earned outside the U.S.?	FOREIGN DEGREES-WIFE

Table 8: Relevant variables that have to be standardized.
Source: Codebooks of PSID Single-Year Family Files interview years 1999 to 2009 see Institute for Social Research, Survey Research Center, University of Michigan, (2007, 2008e, 2008f, 2010b, 2010c, 2011c).

Where CKPT denotes that this variable is the answer to a checkpoint question within the questionnaire, WTR denotes whether, FU denotes Family Unit, HS denotes high school, HD denotes head of Family, HGHST denotes highest, YR denotes year, Coll denotes College, RECD denotes recorded.[92]

[92] See for PSID Single-Year Family Files waves 1999 to 2009 Institute for Social Research, Survey Research Center, University of Michigan (2007, 2008e, 2008f, 2010c, 2010d, 2011c).

2.3.3.2.3 SPSS-Implementation of the solutions outlined in Part B, Section 2.3.3.2.1[93]

This section gives the technical SPSS-implementation of the solutions outlined in Part B, Section 2.3.3.2.1.

2.3.3.2.3.1 Uniform coding scheme and recoding standardized variables

To solve the first problem outlined in Part B, Section 2.3.3.2.1, i.e., different answer possibilities to same question in different waves, a **uniform coding scheme** must be developed that is applied to all waves. This means, for each variable relevant answers must be defined and coded with the same numbers across all waves. This is done by defining value labels. **Value labels** are short explanations of the meaning of a certain parameter value of a variable. For all variables defined, parameter values with their value labels are given below where “Missing” denotes the default value, which means a straight forward answer to the question is not available. Value labels on Wife’s education are the same and are therefore not mentioned separately.

Question: “Whether Family Unit has a new Head in this year.” Variable Label : “Standardized: CKPT: WTR NEW HEAD IN FU”	
Parameter Values	**Value Labels**
1	“New Head”
5	“Same Head”
9999	“Missing”

Question: “Where did you [HEAD] receive your education?” Variable Label : "Standardized: WTR HEAD EDUCATED IN US"	
Parameter Values	**Value Labels**
1	"US only"
2	"Outside US only"
3	"Both US and outside"
5	“No education”
9999	“Missing”

[93] The recoding for each wave can be found in formulas in the SPSS Syntax files as given in footnote 78.

Question: "Did you [HEAD] graduate from high school, get a GED, or neither?" Variable Label : "Standardized: WTR GRADUATED HS-HD"	
Parameter Values	**Value Labels**
1	"Graduated from high school"
2	"Get GED"
3	"Neither"
9999	"Missing; see also STD40573 and ER40574"

Question: "How many grades of school did you [HEAD] finish?" Variable Label : "Standardized: GRADE OF SCHOOL FINISHED-HD"	
Parameter Values	**Value Labels**
1	"Finished first grade"
2	"Finished second grade"
3	"Finished third grade"
4	"Finished fourth grade"
5	"Finished fifth grade"
6	"Finished sixth grade"
7	"Finished seventh grade"
8	"Finished eighth grade"
9	"Finished ninth grade"
10	"Finished tenth grade"
11	"Finished eleventh grade"
9999	"Missing"

Question: "Did you [HEAD] attend college?" Variable Label : "Standardized: WTR ATTENDED COLLEGE-HD"	
Parameter Values	**Value Labels**
1	"Yes"
5	"No"
9999	"Missing"

Question: "What is the highest year of college you [HEAD] have completed?" Variable Label : "Standardized: HGHST YR COLL COMPLETED-HD"	
Parameter Values	**Value Labels**
1	"Completed 1 year"
2	"Completed 2 years"
3	"Completed 3 years"
4	"Completed 4 years"
5	"Completed 5 years or more"
9999	"Missing"

Question: “Did you [HEAD] receive a college degree?” Variable Label : "Standardized: WTR RECD COLLEGE DEGREE-HD"	
Parameter Values	**Value Labels**
1	“Yes”
5	“No”
9999	“Missing”

Question: “What is the highest college degree you [HEAD] have received?” Variable Label : "Standardized: HGHST COLLEGE DEGREE RECD-HD"	
Parameter Values	**Value Labels**
1	"Associate of Arts"
2	"Bachelor of Arts/Science/Letter; BA; BS"
3	"Master of Arts/Science; MA; MS; MBA"
4	"Doctorate, Ph.D. (excepts codes 5 and 6)"
5	"LLB; JD (law degrees)"
6	"MD; DDS; DVM; DO (medical degrees)"
9	"Other"
9999	"Missing"

Question: “How many years of school did you [HEAD] complete outside of the U.S.?” Variable Label : "Standardized: YRS FOREIGN EDUCATION-HD"	
Parameter Values	**Value Labels**
1	"Completed 1 year"
2	"Completed 2 years"
3	"Completed 3 years"
4	"Completed 4 years"
5	"Completed 5 years"
6	"Completed 6 years"
7	"Completed 7 years"
8	"Completed 8 years"
9	"Completed 9 years"
10	"Completed 10 years"
11	"Completed 11 years"
12	"Completed 12 years"
13	"Completed 13 years"
14	"Completed 14 years"
15	"Completed 15 years"
16	"Completed 16 years"
17	"Completed 17 years"
18	"Completed 18 years"
19	"Completed 19 years"
20	"Completed 20 years"
21	"Completed 21 years"

22	"Completed 22 years"
23	"Completed 23 years"
24	"Completed 24 years"
25	"Completed 25 years or more"
9999	“Missing"

Question: “What was the highest degree or certificate you [HEAD] earned outside the U.S.?” Variable Label : "Standardized: FOREIGN DEGREES-HEAD"	
Parameter Values	**Value Labels**
1	"Completed grammar/elementary/primary school but no secondary or high school"
2	"Started secondary or high school but did not finish"
3	"Secondary or high school diploma"
4	"Associate's degree/teaching license"
5	"Bachelor of Arts/Science/Letter; BA; BS"
6	"Master of Arts/Science; MA; MS; MBA"
7	"Doctorate, Ph.D."
9999	„Missing"

Table 9: Standardized value labels for all variables on Head’s education in the Single-Year Family Files. Standardized variables on Wife’s education are labeled the same.
Source: Own definitions and tables.

The second step of standardization is to **recode standardized variables** to conform to the uniform coding scheme. Recoding algorithms may vary from wave to wave as the original coding schemes vary, too.

2.3.3.2.3.2 Characterizing standardized education variables

The technical characterization of a variable includes its variable name, the scale, the definition of parameter values and their numeric format and so on. To solve the second problem outlined in Part B, Section 2.3.3.2.1 (i.e., different codes for same answers in different waves), standardized variables are created as duplicates of their original counterparts. Therefore, initially standardized variables have the same parameter values as the original ones. Later on they will be recoded according to the uniform coding schemes outlined above in Table 9, p. 58. In contrast to overwriting the existing variables creating new ones has the advantage of not losing the information in the original variables. The standardized variables are added to the raw data files. Their characterization of a variable includes its name, label, scale, format and parameter values. This is the first step performed in the SPSS Syntax and will be explained below in detail.

Variable Name

Variables in the Single-Year Family Files from 1989 to 1993 are named V16301 to V23363, variables in the Single-Year Family Files from 1994 to 2007 are named ER2002 to ER41069 - all numbered consecutively. In order to distinguish standardized variables from original variables and include a clear link to their original counterparty at the same time, standardized variables are named "STD" followed by the number of the original variable. E.g., the original variable ER23388 from the 2003 wave is standardized as STD23388.

A closer look at variable's names from 1989 to 2007 reveals that some numbers repeat themselves over the years. Thus, numbers cannot uniquely identify a variable. For example, variables in the 1993 wave are names V21601 to V23363. The same numbers occur in the variable names of the 2003 wave ER21001 to ER24180. Therefore, standardized variables for waves before 1994 are named "STD" followed by an additional "V". This assures a unique identification of all standardized variables and a clear link to their original counterparts at the same time. Standardized variables are called STD-variables and are named as follows:

1994 – 2007 wave:	The original variable ER23388 is standardized as STD23388.
1993 and earlier waves:	The original variable V23245 is standardized as STDV23245.

Variable Label

Variable labels are short explanations of the meaning of a variable. In contrast to the variable name, they can contain several words and/or numbers. Standardized variables are labeled like their original counterparts with "Standardized:" set in front. For example, the original label "CKPT: WTR NEW HEAD IN FU" turns into "Standardized: CKPT: WTR NEW HEAD IN FU". This again assures, that standardized and original variable are comparable.

Variable Scale and Format

Standardized variables are assigned the same scale and format as the original variables.

Variable Parameter Values

Initially, all standardized variables have the same parameter values as the original variables, as they are duplicates. In order to make variables referring to the same questions comparable across different waves, the initial values must be recoded.

2.3.3.2.4 Data consistency check

2.3.3.2.4.1 Problem, solution, and algorithm

The following inconsistency problems only relate to variables on education. Socio-economic characteristics gender and age are not affected.

Problem

A close look at the various education variables reveals that in some records information in one variable contradicts the information given in another variable. An example: 3 years of college completed are recorded in contradiction to another variable indicating the person did not attend college at all. If the educational attainment of a migrant cannot be clearly identified, his income parameters cannot be estimated and he must be excluded from the empirical analysis.

The problem applies to U.S. education as well as the foreign education variables. The latter will be addressed only in the next subsection because it is almost impossible to take into account different education systems, data errors, and the interaction with U.S. education at the same time. The actual numbers of inconsistent records can only be determined after checking for it.

Solution

All standardized U.S. education variables are checked for inconsistent statements with all other standardized education variables (see Figure 5, p. 63). The consistency check is performed after moving to standardized variables because it is easier to check standardized variables with their homogenous coding schemes rather than raw data variables. This way the same algorithm can be applied to all Single-Year Family Files.

The most desirable solution when inconsistencies are detected would be to delete the respective records. In this analysis it is preferred, however, to correct inconsistencies if possible. Deleting inconsistent records would reduce the relevant sample size of economic migrants to a level where a statistically reliable estimation would be difficult. The total number of records within each wave and the percentage of inconsistent records for each variable and wave for Head's records are given below in Table 10, p. 61.[94] Fortunately, all inconsistent records can be corrected.

[94] See exemplary for wave 2009 SPSS-Output files "fam2009er_data Standardize (1-4).spv" and "fam2009er_data Standardize (5).spv" in folder „\Panel Study of Income Dynamics\fam2009er US Haushaltspanel (PSID)".

	Year						
	2009	2007	2005	2003	2001	1999	1997
Total records	8690	8289	8002	7805	7401	6997	6747
SPSS Variable Label							
WTR HEAD EDUCATED IN US	0.7%	43%[95]	51%	55%	3%	3%	3%
WTR GRADUATED HS-HD	0%	0%	0%	0%	0%	0%	0%
WTR ATTENDED COLLEGE-HD	0%	0%	0%	0%	4%	0.2%	1%
WTR RECD COLLEGE DEGREE-HD	0%	0%	0%	0%	0%	0%	0%
HGHST COLLEGE DEGREE RECD-HD	0.1%	0.1%[96]	0.2%	0.3%	0.1%	0.02%	0.01%

	Year							
	1996	1995	1994	1993	1992	1991	1990	1989
Total records	8511	10401	10764	9977	9829	9363	9371	7114
SPSS Variable Label								
WTR HEAD EDUCATED IN US	-	-	-	-	-	-	-	-
WTR GRADUATED HS-HD	0%	0%	0%	0%	0%	0%	0%	0%
WTR ATTENDED COLLEGE-HD	0%	0.01%	0%	0%	0%	0%	0%	0%
WTR RECD COLLEGE DEGREE-HD	0%	0%	0%	0%	0%	0%	0%	0%
HGHST COLLEGE DEGREE RECD-HD	0.01%	0%	0.04%	0.6%	0.5%	0.5%	0.5%	0.3%

Table 10: Percentage of records that had to be corrected for each wave and variable on Head's education. Source: Own calculation.

The high percentages of corrections in 2003 to 2007 are uncritical because they relate to the question “whether Head was educated in the U.S.” This question only summarizes information given in other variables and does not contain new information itself. It can therefore easily be corrected by information contained in other variables. For example, in 2007 of all 8,289 records of Heads included in the file, 3,592 records had to be corrected of which 20 records were originally coded as “U.S. only”, 8 were coded “outside U.S. only”, 9 were coded “both in U.S. and outside”, while the vast majority of 3,555 records was coded “education data last collected before 1997”. In the latter case I recode variables to “missing” which means prior waves are searched through for education information. This can be done because people are asked about their education when they first appear in the Panel Study of Income Dynamics and this information is only updated if their education level changes compared to the previous wave.

For Wives the pattern of corrected records looks alike and is also petty low except for the uncritical question of “whether Head was educated in the U.S.

[95] Of all together 8289 records in 2009, 3592 records had to be corrected of which 20 records were originally coded as “U.S. only”; 8 “outside U.S. only”; 9 “both in U.S. and outside” and 3555 “education data last collected before 1997”.

[96] 0,1% corresponds to a total of 16 records which were all originally coded as “don't know” or “not available” in the raw data.

Algorithm

First, inconsistencies are detected by checking variables in the same order they are asked in the questionnaire (see Figure 5, p. 63). Each answer given is checked in relation to all questions asked afterwards.[97] This way, the same pair of variables will not be checked twice. Second, inconsistencies can be corrected following the logic of the questionnaire where some questions will only be asked if certain answered were given before. To give a better understanding of this logic, the flow chart in Figure 5, p. 63 gives a stylized structure of the questionnaire concerning Head's education. In comparison to the official questionnaire[98], questions irrelevant to this study[99] are left out and answers are given as defined by the standardized variables describes above.

The flow chart shows that, for example, questions about U.S. education like graduating from high school are only asked if the person has previously answered he actually received any education in the U.S. If the person said he only received education outside the U.S., he is only asked about the number of foreign school years and the foreign degree achieved, but not about high school graduation in the U.S.

Following the logic of the questionnaire, inconsistencies can be corrected as follows. For example, if a record includes parameter values other than "missing" for questions on U.S. education even though it is indicated that he did not receive education in the U.S., the latter variable is recoded to "received education in the United States only" or "received education both in the U.S. and abroad", respectively – depending on the parameter values on foreign education. Another example: If a bachelor's degree is recorded even though it is indicated that the person did not attend college, a bachelor's degree is assumed. The reason is that following the logic of the questionnaire the person would not have been asked about his college degree (here Bachelor) if he said he has never been in college. Therefore, the question of whether the person attended college is recoded to "yes". All other pairs of variables are checked using the same logic. For a more detailed code please refer to the SPSS-implementation section.

[97] Except for the question „What is the highest college degree you have received", variable labeled „HGHST COLLEGE DEGREE RECD-HD". Here also questions asked before are considered.
[98] See for example, Institute for Social Research, Survey Research Center, University of Michigan (2010d), pp. 137 – 142 or any questionnaire of another wave, section on Head's and Wife's background and education.
[99] Questions about the exact date when degrees were awarded or school was finished are left out.

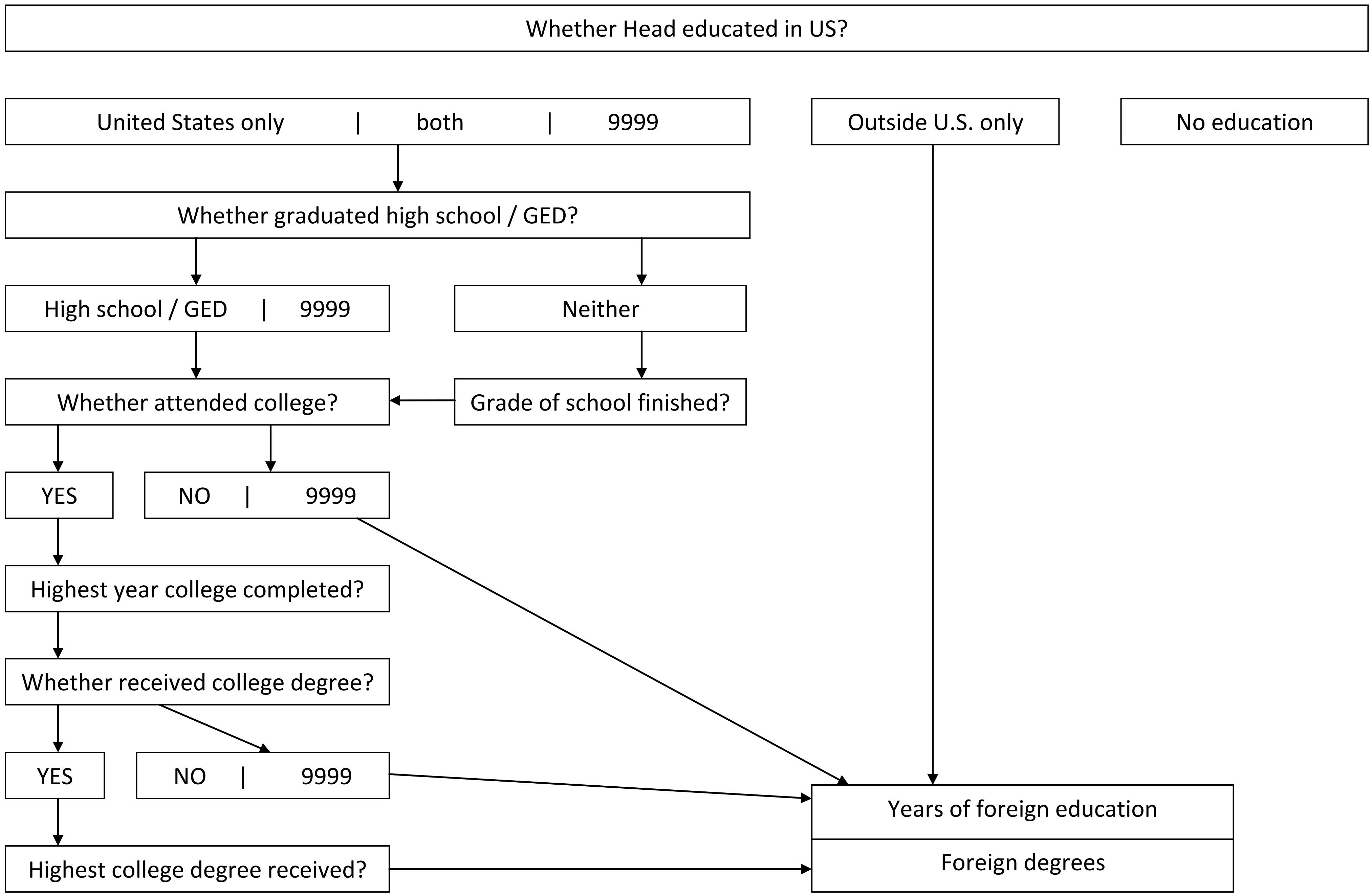


Figure 5: Structure of education questionnaire.
Source: Own illustration.

2.3.3.2.4.2 SPSS-implementation

For each U.S. education variable checked, a Flag-variable is created which indicates whether there is an inconsistency and whether it could have been corrected or not. Flag variables are auxiliary variables with an initial value of zero and value labels defined as follows:

VALUE LABELS Flag-Variables

0	NO inconsistency
1	(automatically) corrected because of inconsistency
5	Inconsistency; check for it manually.

In detail, the consistency check is performed with the following algorithm. The percentage of records corrected is given in Table 10, p. 61.

Variable: "Whether Head educated in the US"

- If indicated that Head received education "only in the U.S." but Head actually received foreign education as well, then the answer is recoded to receive education "both in U.S. and outside".
- If indicated that Head received education "only outside the U.S." but Head actually received U.S education as well, then the answer is recoded to receive education "both in U.S. and outside".
- If indicated that Head received education "both in U.S. and outside" but Head actually did not received education in the U.S. (outside the U.S), then the answer is recoded to receive education "outside U.S. only" ("U.S. only").
- If indicated that Head received "no education" but Head actually did receive education in the U.S. and (or) outside the U.S, then the answer is recoded to receive education "both U.S. and outside" ("outside U.S. only" respectively "U.S. only").
- If the variable on "whether Head educated in the US" has a missing value even though education is reported, it is recoded according to the reported education. In detail this means: (i) If only U.S. education is reported, while variables on foreign education have missing values, the person must be classified as "educated only in the U.S.". (ii) If only foreign education is reported while variables on U.S. education have missing values, the person must be classified as "educated only outside the U.S.". (iii) If variables on U.S. as well as foreign education have parameter values other than "missing", the person must be classified as "educated both in the U.S. and outside". These three cases are the most frequent types of recoding necessary. The reason is that in the raw data the most frequent answer is "education data last collected before 1997" which is then

recoded to "missing" during the standardization process because it is not a straight forward answer to the question of whether Head was educated in the U.S, abroad or both. Still, these records usually contain education information. By looking at the data given, one can conclude where Head was educated and correct the variable parameters.

- It is important to discriminate between those with actually missing education data (because they were not asked these questions) and those, who did not receive any education and are hence coded with default values (interpreted as missing values, too). This difference can be seen in the data because if somebody was asked about education and did not receive it, questions about having a high school diploma or whether attended college are answered with "no". Consequently, someone with parameter values unequal to default values even if he did not receive education is categorized as "no education" received.

Variable: "Whether graduated from high school, get a GED or neither"

If indicated that Head "graduated from high school" or "got a GED" but at the same time grade of school finished is given, this indicates an inconsistency that cannot be corrected automatically as the questionnaire indicates that the latter question will only be asked if Head neither graduated from high school nor got a GED. The corresponding Flag-variable is set to 5; fortunately, this case has not occurred.

Variable: "Grades of school finished"

Here no inconsistency is possible as further questions are about college and there might be many ways to access college.

Variable: "Whether attended college"

If indicated that Head "did not attend" college, but at the same time claims to have completed at least one year in college or to have a degree, the answer is recoded to "attended college".

Variable: "Highest year college completed"

Here no inconsistency is possible as the years completed in college do not necessarily indicate any degree.

Variable: "Whether recorded a college degree"

If indicated that a college degree was "not recorded" but at the same time a certain degree is claimed, the answer is recoded to "college degree awarded".

Variable: "Highest college degree awarded"

If no college degree is mentioned even though the person claims to have one, the answer will be recoded to college degree = "other".

2.3.3.3 A single variable on Head's/Wife's education in the family Files[100]

2.3.3.3.1 Problem, solution, and variable definition

Problem

Each wave of Single-Year Family Files includes many variables concerning Head's education, namely up to 10 variables in each wave. These variables give pretty detailed education information, while it is only the highest educational attainment that is needed for the analysis at hand. This means only a single variable is needed that indicates the highest educational attainment. The same problem holds for Wife's education. As the procedure is the same for Heads and Wives, it is only described for Heads here.

Solution

A new variable is created that summarizes all 10 standardized education variables of Head available in the Single-Year Family Files to indicate only Head's highest educational attainment. It is named STD_HeadEDU07[101] for the 2007 wave, beginning with "STD" as it is a standardized variable. Its coding scheme is as follows:

0	no education
1	finished 8th grade or less
2	9th to 11th grade, but no high school graduate or GED
3	High school graduate/GED
4	Associate's degree
5	Bachelor's degree
6	Master/Profession degree
7	Doctorate, Ph.D.
1111	Error: variable couldn't be coded
9999	Missing.

[100] Corresponding SPSS-Syntax files of, for example, wave 2009 can be found under „\Panel Study of Income Dynamics\fam2009er US Haushaltspanel (PSID)\fam2009er_data Standardize (1-4).sps" concerning Head's education and „\Panel Study of Income Dynamics\fam2009er US Haushaltspanel (PSID)\fam2009er_data Standardize (5).sps" concerning Wife's education. Files and folders for all other waves are named analogously.
[101] The corresponding variables for wives are called STD_WifeEDU07 for 2007.

2.3.3.3.2 Algorithm

In general, the highest level of education someone achieved in any country is determined and coded as mentioned above. In most cases this can easily be done with common sense, e.g., when someone reports a high school diploma and a bachelor's degree, his highest education level obviously is a bachelor's degree. Still the different structure and coding algorithm of education variables on U.S. and foreign education raises multiple questions as the following example shows.

- VALUE LABELS of **U.S. degrees** as defined by the Panel Study of Income Dynamics[102]
  ("What is the highest college degree you have received?")
  1 Associate of Arts
  2 Bachelor of Arts/Science/Letter; BA; BS
  3 Master of Arts/Science; MA; MS; MBA
  4 Doctorate, Ph.D. (except codes 5 and 6)
  5 LLB; JD (law degrees)
  6 MD; DDS; DVM; DO (medical degrees)
  9 Other
  college level
  9999 Missing.

---

[102] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 1362.

- VALUE LABELS of **foreign degrees** as defined by the Panel Study of Income Dynamics [103]

  ("What was the highest degree or certificate you earned outside the U.S.?")

  1 Completed grammar/elementary[104]/primary school but no secondary or high school

  2 Started secondary or high school but did not finish

  3 Secondary or high school diploma

  4 Associate's degree/teaching license

  5 Bachelor of Arts/Science/Letter; BA; BS

  6 Master of Arts/Science; MA; MS; MBA

  7 Doctorate, Ph.D.

  (4–7: College level)

  9999 Missing.

Where BA denotes Bachelor of Arts, BS denotes Bachelor of Science, MA denotes Master of Arts, MS denotes Master of Science, MBA denotes Master of Business Administration, Ph.D. denotes Doctor of Philosophy, LLB denotes Bachelor of Laws, JD denotes Juris Doctor, MD denotes Doctor of Medicine, DDS denotes Doctor of Dental Surgery, DVM denotes Doctor of Veterinary Medicine, DO denotes Doctor of Osteopathic Medicine.

Apparently, huge differences in the variable structure exist which need further explanation. First, the numbers associated with one and the same answer are not the same for U.S. and foreign degrees. Second, the answer possibilities are different. These problems are outlined in detail and finally solved in the following subsections.

2.3.3.3.2.1 Non-comparability of foreign and U.S. education

Problem 1

Foreign and U.S. degrees are not comparable below associate's degree. This raises the question of how to make lower levels of foreign education comparable to U.S. degrees.

Problem 2

There are potential inconsistencies between the two variables on foreign education: school years completed and degrees awarded. Because of different education systems around the world, the number of school years completed abroad may not correspond to the degree recorded if the U.S.

[103] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 1364.

[104] Note that while elementary schools in the U.S. typically educate children from first grade through fourth grade, sometimes up to the eights grade depending on the regional school system (see U.S. Department of Education, International Affairs Office (2008a), Figure 1), grammar schools in England are secondary schools (see National Grammar Schools Association [No date]) which educates children being 11 and older (see Gordon, Peter, and Denis Lawton (2003), p.3). This already points out the difficulty in comparing foreign education systems.

school system is the benchmark. For example, in the United States a master's degree is usually awarded after 5 to 6 years of college. In contrast, the German diploma - which can be assumed to be a Master equivalent - is usually awarded after 4 years of academic studies. The problem becomes more apparent when looking at the average number of school years completed and the corresponding degrees reported in the data. Table 11, p. 69 gives the mean of "foreign school year completed" reported in the American Community Survey between 1999 and 2009[105] in comparison to the U.S. school system as benchmark.

Degree	Mean of "foreign school years completed" reported for a certain degree	US benchmark: years of school to be taken to earn a certain degree in the US
Grammar school	6.1	1 – 8 grade
Secondary not finished	9.0	9 – 11 grade
High school	11.4	12 grade
Associate	14.2	14 grade
Bachelor	14.9	16 grade
Master	14.6[106]	17/18 grade
Doctorate	16.1	19+ grade

Table 11: U.S. education system versus mean years of school reported for foreign degrees.
Source: Own calculations based on the Panel Study of Income Dynamics public release family data files 1999 to 2009 without wave 2003 where foreign degrees are not reported; The corresponding SPSS-Syntax can be found under „\Panel Study of Income Dynamics\average school years outside - foreign degrees.sps"; For the U.S. schooling system see U.S. Department of Education (1995), Figure 1.

It is not clear which variable on foreign education should be used to classify individual's foreign educational attainment for two reasons: First, school systems vary a lot across the world. For example, secondary school usually starts with 9th grade in the U.S., while in Germany it starts at 5th grade. Second, the term "school" has no unique definition as the educational system varies across the world. Therefore, it cannot be excluded that, for example, any type of preschool is considered as school. Hence, using school years to classify individual's education might be biased.

Clearly, both variables have an inherent risk of bias. Even worse, because of these potential biases they might contradict each other, for example, when somebody claims to have a degree equivalent

[105]Family data 2003 is not included because the variables on foreign degrees do not exist for Head and Wife. Even though the respective question is given in the original questionnaire (see Institute for Social Research, Survey Research Center, University of Michigan [No date c], p. 122, Question K54L61) it is left out in the codebook (see Institute for Social Research, Survey Research Center, University of Michigan (2010b), p. 691 for Head and p. 652 for Wife) and the variable itself is missing in the raw data set.
[106] Usually, the average number of school years to be completed to earn a Master degree should be higher than for a Bachelor degree. The numbers given here are correct. The lower number of average school years to get a Master degree compared to the mean years of school to get a Bachelor degree could be due to divergent school systems in other countries.

to high school level after only 8 years of schooling abroad, and no other education in the U.S. This raises the question of what can be done to make foreign education below college level comparable to U.S. education.

Solution

Both problems can be solved at the same time if an auxiliary variable on foreign education is created (STD_HeadforeignEDU07[107]) that has the same coding schemes as the final education variable STD_HeadEDU07:

1 no foreign education
2 finished 8th grade or less
3 9th to 11th grade, but no high school graduate or GED
4 High school graduate/GED
5 Associate's degree
6 Bachelor's degree
7 Master/Profession degree
8 Doctorate, Ph.D.
9 1111 Error: variable couldn't be coded.

The starting value of STD_HeadforeignEDU is set to "1111" and will be recoded if possible.

The coding scheme of the final standardized education variables reflects the U.S. education system. By using the same coding scheme on foreign and U.S. education, both are categorized in the same way. This means, they are comparable even below college level (Problem 1). The idea is that as long as foreign degrees are reported, they are taken for granted.[108] Only if this information is missing, the number of school years completed is transformed to equivalent degrees taking the U.S. school system as benchmark (Problem 2). For example, if someone reports he completed grammar school abroad this is classified as U.S. equivalent "finished 8th grade or less", no matter how many years of foreign education are reported. Only if information on foreign degrees is missing, the reported school years taken abroad are transformed into the U.S. equivalents taking the U.S. education system as

[107] The corresponding variable on wives is named STD_WifeForeignEDU07.
[108] Foreign degrees are systematically missing for waves 1997 and 2003. In the 1997 wave variables on foreign degrees for Head and Wife exist, but all values are set to „inapplicable". In the Single-Year Family File of 2003 the variables on foreign degrees do not exist for Head and Wife. Even though the respective question is given in the original questionnaire (see Institute for Social Research, Survey Research Center, University of Michigan [No date c], p. 122, Question K54L61) it is left out in the codebook (see Institute for Social Research, Survey Research Center, University of Michigan (2010b), p. 691 for Head and p. 652 for Wife) and the variable itself is missing in the raw data set.

benchmark. For example, everything up to 8 years of schooling is classified as “finished 8th grade or less”. The algorithm given below is the same for Wife’s foreign education and all waves.

Table 12, p. 71 gives the coding algorithm if foreign degrees are reported. In this case the number of foreign school years taken is not of interest anymore. Foreign degrees are then transformed to U.S. equivalent degrees like given below.

Foreign degree	**Foreign school years completed**	**U.S. equivalent education gained abroad (STD_HeadForeignEDU07)**
Completed grammar/elementary/primary school but no secondary or high school	Not of interest	Finished 8th grade or less
Started secondary or high school but did not finish	Not of interest	9th to 11th grade, but no high school graduate or GED
Secondary or high school diploma	Not of interest	High school graduate/GED
Associate's degree/teaching license	Not of interest	Associate’s degree
Bachelor of Arts/Science/Letter; BA; BS	Not of interest	Bachelor’s degree
Master of Arts/Science; MA; MS; MBA	Not of interest	Master’s degree
Doctorate, Ph.D.	Not of interest	Doctorate, Ph.D.

Table 12: Coding algorithm of the variable “STD_HeadForeignEDU07” on foreign education if the original variable “foreign degree” is not missing.
Source: Own illustration.

Table 13, p. 71 gives the coding algorithm if foreign degrees are missing. In this case the number of foreign school years are transformed to equivalent U.S. degrees using the U.S. school system as benchmark. The detailed recoding is given below.

Foreign degree	**Foreign school years completed**	**U.S. equivalent education gained abroad (STD_HeadForeignEDU07)**
Missing	1 to 8 years	Finished 8th grade or less
Missing	9 to 11 years	9th to 11th grade, but no high school graduate or GED
Missing	12 to 13 years	High school graduate/GED
Missing	14 to 15 years	Associate’s degree
Missing	16 to 17 years	Bachelor’s degree
Missing	18 years and more	Master’s degree

Table 13: Coding algorithm of the variable “STD_HeadForeignEDU07” on foreign education if the original variable “foreign degree” has missing values.
Source: Own illustration.

After coding the auxiliary variable STD_HeadForeignEDU07, it can be compared to the U.S. education of Head and the highest educational attainment in any country is defined as the (overall) education level of Head, given by the variable STD_HeadEDU07.

2.3.3.3.2.2 Medical/Law/Honorary degrees

Problem

College level degrees awarded in the U.S. have additional categories for medical and law degrees as outlined in Part B, Section 2.3.3.3.2. This poses the question of how to subsume medical and law degrees under the Bachelor, Master and Doctorate categories of the uniform coding scheme defined for the standardized Head's education variable (see Part B, Section 2.3.3.3.1).

Solution

In my dissertation the educational attainment of people is meant to be a determining factor of income. This must be kept in mind when medical degrees, law degrees, and honorary degrees are classified under the categories of the standardized education variable.

Medical degrees

Medical degrees are classified as "Doctorate" equivalent because medical degrees in the U.S. are First-Professional Degrees that require "a total of at least 6 academic years of college work to complete the degree program".[109] Students wanting to enter medical school must have completed at least 2 years of college before[110] and are usually asked to have passed Medical College Admission Tests.[111] According to the answer possibilities of the original questionnaire[112] this applies to: Medical Doctor (M.D.); Doctor of Dental Surgery (D.D.S.); Doctor of Veterinary Medicine (D.V.M.); Doctor of Osteopathic Medicine (D.O.).[113]

Law degrees

Bachelor of Laws (L.L.B.) and the Juris Doctor (J.D.) are the lowest level of law degrees available which are usually followed by the Master of Laws (LL.M.) on the graduate level.[114] The highest law degree is the Doctor of Juridical Science (S.J.D.) which is primarily aimed at training professors, legal scientists, and other scholars in law.[115]

---

[109] See U.S. Department of Education, International Affairs Office (2008b).
[110] See U.S. Department of Education, International Affairs Office (2008b).
[111] See Association of American Medical Colleges (2012).
[112] The original coding was already presented in Part B, Section 2.3.3.3.2.
[113] See U.S. Department of Education, International Affairs Office (2008b).
[114] See Harvard Law School (2012b).
[115] See Harvard Law School (2012a).

Since higher level law degrees are not mentioned explicitly in the questionnaire, it is plausible to assume that they are subsumed under Master (LL.M.) or Doctorate (J.S.D. or S.J.D.). Therefore, L.L.B. and J.D. are classified as Bachelor equivalent.

Honorary degrees

The WordNet, a lexical database of Princeton University, defines an honorary degree as a degree to honor the recipients.[116] It must not be mistaken for an ordinary academic degree that is awarded with honors, e.g., an Honors bachelor's degree.[117] In the Panel Study of Income Dynamics honorary degrees seems to refer to an honorary doctorate which usually includes the letters "h.c." after the award to indicate the status. In my analysis it is not classified as any educational attainment for three reasons: First, the reason for such a degree could be almost everything from academic contributions to appraisal of a political or social attitude. Second, it can be assumed that people with an academic title will report this rather than an honorary degree. Third, there is not a single Head in the Family Files from 1989 to 2009 who reports an honorary degree.

2.3.3.3.2.3 Imprecise data: "other" U.S. college degrees

Problem

A degree classified as "other" could be any degree, license or certificate. This problem only refers to U.S. degrees as the option is non-existing for foreign degrees.

Solution

Combined with the information about years of college completed the following coding scheme is applied:

- "1 year of college completed" is set to "high school graduate": Someone who reports a degree after only one year of college may have a degree below the lowest level defined (Associate) or does not have one at all (data error). Therefore he is recorded as high school graduate as this is the next lowest level.

[116] See Princeton University (2012).
[117] See U.S. Department of Education, International Affairs Office (2008e).

- “2-4 years of college completed” is set to “Associate’s degree”: Any degree awarded after 2 to 4 years of college is defined to be equivalent to the lowest level mentioned in the data. This is plausible as an associate’s degree usually takes at least 2 years of college.[118] For 3 and 4 years of college no higher degree is defined as a bachelor’s degree usually takes 4 years and more of college.
- “5 and more years completed” is set to “Bachelor”: Any degree awarded after five or more years of college is defined as Bachelor equivalent as the regular study period is 4-5 years.

Higher numbers of years are not defined as Master level as this level is rather advanced and therefore the potential bias if a master’s degree is assigned falsely might be great. In the Single-Year Family Files from 1989 to 2009 a college degree classified as “other” for Heads or Wives is only reported for less than 1% of the families.[119]

2.3.3.3.2.4 College attendance without degree

Problem

If people are reported to have attended college and even completed certain years of college but do not have a reported degree, it is unclear whether they just failed to earn a degree or the variable on college degree has a missing value.

Solution

In order not to overestimate the level of education but still account for the years completed at college an even more cautious recoding than in the sections above is performed.

- “1-3 years of college completed” is set to “high school graduate”: In this case recoding must be exercised with caution. Therefore, people with 1-3 years of college completed are classified as high school graduates.

[118] See U.S. Education Department (1995), Figure 1.
[119] See Codebooks of PSDI Single-Year Family Files for interviewing years 1989 to 2009: Institute for Social Research, Survey Research Center, University of Michigan (1992a, 1992b, 1995a, 1995b, 1995c, 2007, 2008a, 2008b, 2008c, 2008d, 2008e, 2008f, 2008g, 2010c, 2010d, 2011c).

- "4 years and more of college completed" is set to "Associate's degree": After 4 years or more of College one might assume that this could be considered equivalent to the lowest college degree associate's degree, which is usually awarded after 2 years.

Here again, higher degrees are not assigned as the risk of getting a bias is too high.

2.3.3.4 Collecting all education information in one single file[120]

2.3.3.4.1 Problem, solution, and graphical overview of data set

Problem

Education variables are spatially scattered across the Cross-Year Individual File and the 15 Single-Year Family Files from 1989 to 2009. Thus, it is impossible to say what type of information is available in which year. Thus a clear definition of education at the time of the move is also impossible. But this information is needed for an empirical analysis of the theoretical migration decision model.

Solution

Education information on all individuals in the Panel Study of Income Dynamics between 1989 and 2009 is collected in one single file. The final file includes (i) the number of school years completed (taken for each wave from the Cross-Year Individual File) and (ii) Head's and Wife's highest academic degree (taken from the 15 Single-Year Family Files). This way a good overview of all education information available for each person in the Panel Study of Income Dynamics from 1989 to 2009 is created. This education overview is added to the Cross-Year Individual File. The following Figure 6, p. 76 gives the stylized structure of the resulting data file.[121]

[120] The corresponding SPSS-Syntax file can be found under "\Panel Study of Income Dynamics\ Collect ALL (Head,Wife;FMembers) EDU (file 1-30).sps". The final SPSS-data file 30_Collected EDU of ALL 89-09.sav can be found in the same folder.

[121] The resulting SPSS data file can be found under "\Panel Study of Income Dynamics\ 30_Collected EDU of ALL 89-09.sav".

Family Interview Numbers 2009 - 1989			Education information 2009				. . .	Education information 1989				Cross-Year Individual File 1968 - 2009	
IntNo09	…	IntNo89	Sequence Number 2009	Head's highest education level from Single-Year Family File 2009	Wife's highest education level from Single-Year Family File 2009	Years of education completed from Cross-Year Individual File 2009	…	Sequence Number 1989	Head's highest education level from Single-Year Family File 1989	Wife's highest education level from Single-Year Family File 1989	Years of education completed from Cross-Year Individual File 1989	…	…
5654	…	2815	1	5	3333	16	…	4	3333	3333	0	…	…
744	…	0	2	3333	2	12	…	0	3333	3333	0	…	…
2133	…	5244	5	3333	3333	0	…	0	3333	3333	0	…	…
…	…	…	…	…	…	…	…	…	…	…	…	…	…

Figure 6: Stylized structure of data file with education information of all years and types of all individuals added to the Cross-Year Individual File.
Source: Own illustration of the SPSS data file 30_Collected EDU of ALL 89-09.sav with exemplary individuals.

Education levels of Heads/Wives that are labeled “3333” indicate that the person is not Head/Wife in that year. Note, that this is not the same as the default value “9999” which means that the person is Head/Wife, but education information is missing.

Examples to read Figure 6, p. 76 may illustrate the data structure: The first person given in Figure 6, p. 76 belongs to the family with the Family Interview Number 5654 in 2009 and 2815 in 1989. The individual is Head in 2009 (Sequence Number equals 1). His education level as given in the Single-Year Family File is 5, meaning his highest degree earned is a bachelor’s degree. As the person is Head, not Wife, the corresponding education variable on Wife is coded “3333”. In the Cross-Year Individual File 16 years of education are reported. In 1989 the same person was neither Head nor Wife in his family but ranked fourth (Sequence Number 4). That could be, for example, the second born child to a married couple. Therefore, Head’s and Wife’ education variables from the Single-Year Family Files are set to “3333”. The Cross-Year Individual Family File reports zero years of completed education in 1989.

The second person in Figure 6, p. 76 is ranked second within her 2009-family (Wife). She is not associated to a family in 1989 (Family Interview Number 1989 equals zero), consequently no education information about her in 1989 is available. The last person in Figure 6, p. 76 is ranked as number 5 in the 2009-family. Since he is neither Head nor Wife, no education information is available from the Single-Year Family Files (code “3333”). The Cross-Year Individual File reports zero years of completed education. All other parameters given for the second and third person in Figure 6, p. 76 can be interpreted analogously to the first person.

2.3.3.4.2 SPSS-implementation

Each person in the Cross-Year Individual File (1968 to 2009) is associated with a family in the Single-Year Family File by his annual Family Interview Number.[122] Furthermore, the personal annual Sequence Number gives the rank of that individual within the family unit at the time of the survey.[123]

Using the annual Family Interview Number, all standardized education variables of Heads and Wives of the corresponding Single-Year Family Files for each wave from 1989 to 2009 are added step by step to the Cross-Year Individual File. In case the individual was not Head/Wife in that year

[122] Here it is important to note that Family Interview Numbers most certainly change from year to year for one and the same family as the annual interview numbers are assigned based on receipts of the interview, i.e., the first interview coming in from field is numbered 1, the second 2, and so on (see Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3.

[123] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 6.

(Sequence Numbers greater than 2) the added education variables of Head/Wife are overwritten with codes “3333”. If this was not done, the variable would suggest that the individual had the education of Head/Wife of his corresponding family in that year. Finally, for a better overview variables are sorted by years as described in Figure 6, p. 76.

2.3.3.5 Making education in the Cross-Year Individual File and the Single-Year Family Files comparable

Problem

The Panel Study of Income Dynamics offers two different types of education information depending on the rank of that person within his family: (i) The Cross-Year Individual File gives the number of years completed at school up to 17 years *for every person*. (ii) The Single-Year Family Files include information about the highest academic degree achieved *only for Heads and Wives*.[124] This raises the question of how to compare the two variables as a uniform definition of education is needed for all family members. This problem is similar to the one outlined in Part B, Section 2.3.3.3.2.1, namely comparability of foreign school years and U.S. education degrees.

Solution

The number of school years completed (given in the Cross-Year Individual File) is transformed to an equivalent academic degree (like in the Single-Year Family Files). To avoid potential biases, the same recoding scheme used to standardize foreign education and U.S. education is applied here (see Table 13, p. 71).

2.3.3.6 Defining education at the time of the move[125]

2.3.3.6.1 Objective and problems remained open

The objective of this section is to finally find an algorithm that clearly defines education at the time of the move. This is afflicted with two problems: First, although education variables of the Single-Year Family Files and all education variables in the Cross-Year Individual File have the same coding scheme, there are still two variables for education for each year of data, namely one from the Single-

---

[124] Education variables on Heads and Wives have already been standardized across all waves of data (see Part B, Section 2.2) and summarized to the highest academic degree in a single variable (see Part B, Section 2.3.3.3).

[125] The corresponding SPSS Syntax file is (exemplary for migrant wave 2007/2009) „\Mover\(1) Mover between Waves\2007_2009\Std_BringForward_MoverEdu_07_09 (7-8).sps”. The resulting SPSS data files (again, exemplary for migrant wave 2007/2009) is „\Mover\(1) Mover between Waves\2007_2009\ 8_all_econ_mover_&EDU_forward.sav”.

Year Family File and one from the Cross-Year Individual File. Their content may contradict each other (see Part B, Section 2.3.3.6.2) although their coding scheme has been made comparable in the foregoing section. Second, it could be that both variables have missing values since education belongs to the category of "background information" which is only reported once when someone first enters the Panel Study of Income Dynamics. The need to bring forward missing education information raises the question of which direction to go, future or past waves (see Part B, Section 2.3.3.6.3). Finally, all these problems must be solved simultaneously by defining a variable on education at the time of the move (see Part B, Section 2.3.3.6.4).

There are different ways to solve these problems, but all of them have certain drawbacks. Therefore, competing solutions with their advantages and disadvantages are discussed in the following subsections. The resulting education variables will be used to run a robustness check and are summarized in Part B, Section 2.3.3.6.4.

2.3.3.6.2 Contradicting education information in the Cross-Year Individual File and the Single-Year Family Files

Problem

Although made comparable, variables in the Cross-Year Individual File and the Single-Year Family Files may contradict each other. This is, for example, the case for Head of the family with the Family Interview Number 8099 in 2009: For him 8 years of schooling are reported in the Individual Cross-Year File, while the Single-Year Family Files report he has finished "9 to 11th grade".

Solution

The problem of contradicting education information of the two sources (Single-Year Family Files and Cross-Year Individual File) can be solved in three different ways:

- **Take consolidated education:**
  The various education variables in the Single-Year Family Files are more detailed and thus have a higher information content than education variables in the Cross-Year Individual File (see Part B, Section 2.1, Problem 2). Consequently, for each year family level information from the Single-Year Family File is preferred whenever available. Only if the person was neither Head nor Wife in that year or no education was reported, his education for that year is set equal to education information from reported in the Cross-Year Individual File.

- **Check Head/Wife of all waves first:**
  All waves of Single-Year Family Files are checked for family level education information first. Only if the person was never Head or Wife in all waves, individual level variables from the Cross-Year Individual File are taken for granted.

- **Take highest education in each wave:**
  In each year the highest education level reported in both sources is taken for granted. This makes sure that education is not underestimated.

2.3.3.6.3 Direction of bringing forward missing education (from future or past)

Problem: Where to search for missing data (future/past waves)?

Education information of both types is not available in all years. In the Single-Year Family Files questions on Head's education are only asked in the year when Head/Wife first became Head/Wife of that family.[126] Furthermore, in the Cross-Year Individual File questions about years of school taken are also not asked every wave.[127] This raises the question of whether to search for missing information in future or past waves.

Solution

In general, education which has been achieved once cannot be lost. This means going back to previous waves to search for education information makes sure that education is not overestimated. On the other hand, it may be the case that education is underestimated this way as individuals may acquire additional education in the meantime and going back sometimes means going back more than 20 years. To mitigate this problem, it may be reasonable to look for more current information in the years after the move as well.

Consequently, there are good arguments for going back as well as going forth when searching for missing data. This leads to two different approaches: "Going back then forth" and "searching centered around the moving year":

[126] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 1338.
[127] See Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 662.

- **“going back then forth”**

  “Bringing forward background information” goes back to the algorithm recommended by the Panel Study of Income Dynamics itself.[128] The idea behind this approach is that education once achieved cannot be lost in the future. Therefore, it is recommended to go back wave by wave and search for relevant information. In my dataset this entails two problems. First, the data structure before 1989 and the related coding scheme of the Panel Study of Income Dynamics is totally different from waves after 1989. Therefore, only variables from 1989 to 2007 have been standardized. Second, going back further than 1989 may not lead to the desired result as education information which has been collected more than 10 years before the actual move takes place may be outdated. Therefore, the following algorithm is chosen: Going back till 1989 and if nothing is found in the past, going forward till 2009. For example, if education in 2003 is missing, it is searched for education information in the following order: waves 2001 – 1999 – 1997 – … back till 1989 – then forth 2005 – 2007 - 2009. This is what is called “going back then forth”.

- “**centered around the moving year”**

  Searching “centered around the moving year” follows the idea that the most recent information available is the best approximation of education at the time of the move. To illustrate the algorithm let’s take the same example as before. If education in 2003 is missing, it is searched in the following order:

  a) check education in 2001 (one wave back),
  b) check education in 2005 (one wave forth),
  c) check education in 1999 (two waves back),
  d) check education in 2007 (two waves forth),
  e) check education in 1997 (three waves back),
  f) check education in 2009 (three waves forth),
  g) as 2009 is the most recent wave, I cannot go further into the future but only back.
  h) check education in 1996 (four waves back),
  i) check education in 1989 (eleven waves back).

[128] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 84.

2.3.3.6.4 Six definitions of education at the time of the move

If the matter of direction (Part B, Section 2.3.3.6.3) is combined with the question about which data should be preferred (Part B, Section 2.3.3.6.2) six possible definitions of education at the time of the move are obtained:

Variable	Algorithm
STD_EduMove_1	Take consolidated education in each year (family-level preferred over individual-level education) and search centered around the moving year for education data.
STD_EduMove_2	Take consolidated education in each year (family-level preferred over individual-level education) and search previous waves first, then future ones for education data.
STD_EduMove_3	Search for Head/Wife education data centered around the moving year. If nothing is found, check individual-level education data centered.
STD_EduMove_4	Search for Head/Wife education data back then forth. If nothing is found, check individual-level education back and forth.
STD_EduMove_5	Take the highest education reported in each wave; search centered around the moving year.
STD_EduMove_6	Take the highest education reported in each wave; go back then forth.

Table 14: Six definition of education at the time of the move.
Source: Own illustration based on own definitions.

Following the rules of the Panel Study of Income Dynamics, everyone is asked at least once about his education. If all education data from the very beginning of the Panel Study of Income Dynamics has default values it could either be a result of missing data or of missing education since there is no question where people can explicitly indicate that they did not receive any education. Consequently, people without education cannot be differentiated from those having missing values.

2.4 Identifying family ties: the Move Context Variable[129]

This section refers to the general Problem 4 as described in Part B, Section 2.1. It once again outlines why family ties are needed, which information is needed, and which information is available in the Panel Study of Income Dynamics. Finally, it is described how information needed is obtained.

[129] In the corresponding SPSS files the Move Context variable is named „STD_MoveIndicator". It is programmed in the SPSS Syntax file „\Mover\(1) Mover between Waves\2007_2009\Identify_Mover_07_09 (1-6).sps" in step 6. The final SPSS data file is "Data\Mover\(1) Mover between Waves\2007_2009\6_all econ_mover_&MoveIndicator_07_09.sav". All folders are given exemplary for migrant wave 2007/09.

2.4.1 Why family ties are important to my study

Family ties are of great importance to my analysis for three reasons: First, the family migration decision model developed in Part A assumes that Head takes the migration decision based on family income possibilities. Therefore, it is important to know who is part of the family at the time of the move. Second, personal relations may interfere with assumed economic migration behavior. For example, it could be that someone reports economic reasons for his move but at the same time he has decided to leave his old partner and children. In this case it could be that the family composition change interferes with his reported (economic) reason to move and even may contradict it. This would bias the risk-attitude derived from the migration decision observed. To be able to run a robustness check with and without families that may have interfering personal relations, a variable on family ties is needed. Third, the migration decisions of Heads may be influenced by the number and type of family members who move with him, i.e., the total number of family members, the number of dependent children, whether toddlers are associated and so on.

2.4.2 Existing PSID-variables on family ties

Although family ties are important to my study, they are not adequately captured by the Panel Study of Income Dynamics. It only includes variables that reveal partial information about family ties. From the Cross-Year Individual File the following variables are useful to my study:

(1) The **Sequence Number** gives the individual's status within a family.[130] It indicates whether someone actually lives with the family, left the family, died, or lives in an institution by the time of the interview.[131] The most important types of institutions mentioned are facilities of the armed forces, health care facilities like hospitals, education facilities like college dormitories[132], or prisons.[133] Furthermore, those living with the family are numbered consecutive from 1 to 20 starting with 1 for Head.[134] The person numbered 2 being Wife or husband of Head[135], or - if Head has no partner - the second person in the family hierarchy.[136]

[130] See Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 654.
[131] See Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 654.
[132] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 79, Paragraph 3.
[133] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 3.
[134] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 6.
[135] See Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 6.
[136] See Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 654.

(2) The variable "**Relation To Head**" [137] contains very detailed information about family relations by blood or by law like being son/daughter, stepson/stepdaughter, son/daughter-in-law, cousin of Head and so on. Furthermore, other relations are also considered like being parents of the cohabiter; being other non-relative including friends; being first year cohabiter or children of him/her and so on.

(1) The **Marital Pairs Indicator** "links pairs of individuals who were married or permanently cohabiting at the time of the (..) interview. Spouses in the first such pair within a family unit each receive a code value of 1 here; the second, a value of 2, and so on."[138]

(2) The variable **Year Individual Born** is self explaining.[139]

2.4.3 Problems and solution of family ties: the Move Context Variable

In general the Move Context Variable should give a detailed insight into who moves with whom, having which kind of personal relation to whom before and after the move. In order to achieve this objective, the following problems must be solved:

Problem 1: Only a static view of family ties is available

The existing PSID-variables on family ties presented in Part B, Section 2.4.2 are available for the wave before and after the move. In contrast to the static information available, it is the family composition at the time of the move that is decisive for the migration decision. As the family composition can change within the period of moving, looking separately only on family ties before or after the move may draw a wrong picture of family ties at the time of the move. An example may illustrate the problem. Assume a family that consists of parents with three children in the wave before the move. After the move parents are divorced so that the father together with one child is reported as family of two, while the mother and another child move in with her new partner and his child. They are then considered a separated family of four (wife, one child of wife, new partner with child). Irrespective of the parents' separation the third child moves in with his new partner now forming his own separate family of two. If the family composition of these people is only considered before or after the move without looking at the changes that took place, a wrong picture will be drawn. In the wave before the move it seems as if a family of five has been moving – which is not the case. On the other hand, if only the wave after the move is considered, one might get the impression that there

[137] See Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 655/656.
[138] Social Research, Survey Research Center, University of Michigan (2010a), p 657.
[139] See Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 657.

are three separate families moving. This is not the case either. Therefore, a dynamic approach must be chosen, taking into account the family composition change - an approach that is not accounted for in the Panel Study of Income Dynamics.

Problem 2: Several variables on personal relations exist

The Panel Study of Income Dynamics includes several variables on family ties (see Part B, Section 2.4.2) but first, they do not link groups of moving people and second, do not give the relation between people in a single variable but in various variables.

Problem 3: Unknown people

Not all individuals surveyed in the Panel Study of Income Dynamics were interviewed before and after the move. Some individuals are first interviewed when migration has already taken place. For these people data from earlier waves is missing. Especially critical is the fact that it is impossible to figure out where they have lived before the move and whether they have actually moved.[140]

Problem 4: Family members not being Head

People who are not Head are not asked about their reason to move. Thus, their migration decision cannot be modeled on an individual basis but only on the family basis if a moving Head is associated with that person. This makes it necessary to consider these individuals separately from others, leaving the opportunity to cancel out families including unknown people.

Solution to Problems 1 to 4

A single variable called **Move Context Variable** is created. It classifies each person into a system of family ties with four different types of moving constellations: single-moves, pair-moves, family-moves and other moves; all of which are further categorized depending on the family ties before and after the move taking up a dynamic approach. Additionally, the Move Context Variable separates people for whom detailed information about the sending country is not available because they are first interviewed after the move. These people are therefore called "**unknown people**". They can either have a relation to movers (e.g., if they are parents, partners, children, other relatives who were not part of the panel study so far) or they do not have any relation to the mover.

[140] Even though a comparison of former and actual state of residence suggests a move, this would be a wrong conclusion in case of an unknown person. The reason is that the Panel Study of Income Dynamics relates new individuals to the family history of the family they live with when they are interviewed for the first time - although the person was not related to that family in the past.

The Move Context Variable is created by taking into account information from the personal dataset (such as the Sequence Number of an individual within his family before and after the move)[141] and the family dataset, which reveals information about all kind of personal relations, such as relation by blood and adoption, step-relations, foster children, son/daughter-in-law.[142] By comparing all variables from the personal and family dataset before and after the move, a pretty detailed picture of family ties incorporating family dynamics can be drawn.

The Move Context Variable makes it possible to analyze the family migration decision and its resulting risk parameters not only based on monetary determinants but also on personal relationships that may interfere with them.

Creating such a differentiated set of parameter values for the Move Context Variable leaves a lot of opportunities for the analysis of the resulting risk-parameters. Especially those movers in the last category can be easily added or excluded from the family.

2.4.4 Categories of the Move Context Variable

The Move Context Variable **intentionally differentiates** between very similar but slightly different family constellations to keep the opportunity of all types of analysis on the resulting risk-parameters. Together with the annual Family Interview Number the Move Context Variable clearly defines the family ties of each migrant and whole families. In detail, the Move Context Variable is defined as follows:

1.) Single-move

A **single-move** indicates that a person moves on his own without members of the old or new family. Considering the migration decision, this means that he is maximizing his own preference function based on his individual income which is the family income at the same time. The single-move includes the following subcategories:

[141] Relevant variables of the Cross-Year Individual Files are describes in detail in Part B, Section 2.4.1.
[142] For a more detailed overview of relations to Head see Institute for Social Research, Survey Research Center, University of Michigan (2010a), p. 655f.

- **VALUE LABLES MoveContext: Single-move**
  1 Single-move - lives alone (1-person family before and after the move)
  2 Single-move - leaves institution and lives alone now
  3 Single-move - leaves partner and lives alone now
  4 Single-move - leaves family but not partner and lives alone now
  5 Single-move - moves in with parents/new partner/children/other relatives (left old partner)
  6 Single-move - moves in with parents/new partner/children/other relatives (lived in institution/alone/as lower level[143] family member before)
  7 Single-move - moves in with unknown person with no relation, who is not partner (left old partner)
  8 Single-move - moves in with unknown person with no relation, who is not partner (lived in institution/alone/as lower level family member before)

It is accounted for all types of family ties before and after the move. Category (1) indicates that the person lived alone before and after the move. Categories (2) to (4) include all cases where the person lives alone after the move but not before. The other categories (5) to (8) include cases where the person moves in with other people. In these cases it is distinguished between new cohabiters with which the person may have any kind of personal relation like parents, a new partner, children or other relatives (categories (5) and (6)) and those cohabiters for which a personal relation to the single-mover is not probable because there is no indication for this in the data (categories (7) and (8)).

Depending on the family ties the mover had before, the move categories (2) to (8) distinguish between leaving the old partner, leaving the old family without leaving the partner, leaving as lower family member and leaving an institution.

2.) Pair-move

A **pair-move** takes place when two people who are indicated to be a pair move together.

- **VALUE LABLES MoveContext: Pair-move**
  20 Pair-move - two-person-family being a pair before and/or after the move (no newborn)
  21 Pair -move - pair with newborn(s)
  22 Pair -move - pair moves in with parents/children/other relatives/ unknown person with relation
  23 Pair -move – pair moves in with unknown person with no relation

[143] L**ower level family members are defined as** Family members who are not Head.

If two people have lived together as a family before the move, move together to a new place, and are indicated to be a pair only after the move, they are assumed to be a pair when they migrated (category (20)). Even if a pair gets a firstborn between two consecutive waves, they are still considered as pair-move (category (21)). This assumption can easily be relaxed later on when newborns (category (40)) are also considered as family members. It is distinguished between pairs moving with and without newborn(s) because becoming parents may interfere with economic reasons for migration reported in the data. In case of category (20) the pair lives together as a two-person family, in case of category (21) the family consists of the pair and their newborn(s). Sometimes a pair moves in with new people who (i) have not been in the panel data before, (ii) are not newborns, and (iii) do not have any other relation to the pair (category (23)). This case is separated from a pair moving in with parents, children or other relative (category (22)) to point out the potentially different type of personal relation to the new cohabiter(s).

3.) Family-move

Family-moves indicate that it is (i) either at least two people who are no pair before or after the move or (ii) more than two people where two of them could be a pair but not necessarily have to.

- **VALUE LABLES MoveContext: Family-move**
  30 Family-move - whole family moves unchanged
  31 Family-move - whole family moves plus newborn(s) before or after move
  32 Family-move - part of family moves, but not leaving partner
  33 Family-move - split-off move, leaving partner

If all family members move, it is further distinguished between families who move in an unchanged constellation (category (30)) and those who get newborn(s) between two waves of the Panel Study of Income Dynamics (category (31)). If only one part of the family moves, it cannot judged which part of the family can be considered to be the "main family"[144] as there is no variable indicating who left whom for what reason. Therefore, it is only made a distinction between families, where partner separate from each other, so called "split-off families"[145] (category (33)) and those where this is not the case (category (32)), e.g., when children leave their parent-family to build their own economically independent family.

[144] „A main family is one that is the source of a splitoff family (a new study family formed by a sample member who moves out and forms his or her own family unit)", Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 79, Paragraph 2.

[145] „A split-off family consists of a person or group of people (…) who moved out from a main family since the prior wave's interview to form a new, economically independent family unit living in a separate housing unit." Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 79, Paragraph 3.

4.) Other Moves

Finally, **other moves** include all people that cannot be subsumed under the categories mentioned above.

- **VALUE LABLES MoveContext: Other Moves**

  40 Newborn in a family/pair- move
  41 Unknown parent/partner/child/other relative somebody moves in with
  42 Unknown person somebody moves in with, no relation
  43 Irrelevant because person left family/died/moved to institution
  44 Irrelevant because no associated Head moves
  9999 Missing.

First, newborns (category (40)) have their own category because they might not be full family members at the time of the move, which means it is not clear whether Head already considered them in the decision problem or not. Defining newborn(s) as an own category separate from other full family members keeps the opportunity of robustness checks with and without them.

Second, unknown people are separated because it is not clear whether and where they moved. By separating them, the opportunity of a robustness check is left open. They are further divided into people that may have a personal relation to the moving family member, such as parents, partner, children or other relatives (category (41)) and those which seem to have no further personal relation (category (42)).

Finally, there are two types of people who are irrelevant for this study because they either moved without the new Head (category (44)) thus making it impossible to figure out their reason to move, or are not followed by the Panel Study of Income Dynamics because they left their own family, died or moved to an institution (category (43)).

The default value is defined as (9999), which means the case cannot be coded according to the above defined standards. This case does not appear in the data.

The categorization is performed by hand, meaning that each family constellation before and after the move is assessed individually and is then classified into one of the above mentioned categories. Note that not all family members moving together must possess the same Move Context category. For example, a pair moves and gets a baby between the two waves. In this case the parents are categorized as Move Context 21, while the newborn is categorized as Move Context 40. Even though they have an identical Family Interview Number, their parameter value of the Move Context Variable differs.[146]

2.5 Unique identification of individuals and families in the PSID[147]

Problem

This section refers to the general Problem 5 as described in Part B, Section 2.1. The final data set includes of all individuals and families respectively who moved between 2000 and 2009. Unfortunately, the raw data set includes only annual identifiers which repeat themselves every year, while a variable that uniquely identifies individual migrants and migrating families is missing. The reasons are as follows:

Family identification

Concerning the identification of families, the reason is that the Family Interview Number, which identifies families in each wave separately, is assigned based on receipts of the interview in a certain year.[148] This means that, for example, the Interview Number 5 (indicating that this family was the 5th in that year to hand in its questionnaire) is assigned again in every survey year. Therefore, it could happen that within a sample including several years of data one and the same Family Interview Number is assigned to two (or more) different families, making it impossible to uniquely identify families.

[146] The final SPSS data file is „\Mover\(1) Mover between Waves\2007_2009\6_all econ_mover_&MoveIndicator_07_09.sav".

[147] All files can be found in the folder „\Mover\(4) Personal and Family Mover IDs" with the corresponding SPSS-Syntax files named „(4a) Add Personal and Family IDs.sps" and "(4b-c) Delete MoverID and Sort by FamilyMoverID". The coding of the family identification numbers is first performed in the excel file "(4a) Creating FamilyMover_ID.xlsm" and is then copied to the SPSS-data files "4a_FamilyMoverID_2000.sav" (exemplary for migrant wave 2000). Later on these numbers are corrected in order to have consecutive numbers after former families 131 and 155 were deleted from the sample because their Heads could not be defined into a socio-economic groups. The correction is documented in the excel file "(4f) Renaming Family and PersonalMoverIDs.xlsm". The coding of the personal identification numbers is performed in the excel file "(4f) Renaming Family and PersonalMoverIDs.xlsm". The final coding of the identification variables for families and individuals can be found in the SSPS-data file "4f_New Family and PersonalIDs.sav".

[148] The Family Interview Number is explained in detail in Part A, Section 3.4.1.2.3 (see also Institute for Social Research, Survey Research Center, University of Michigan (2011a), Question 3).

Individual identification

Concerning the identification of individuals, a combination of Family Interview Number and Sequence Number[149] uniquely identifies a person only in a single year. Once migrants of different waves are surveyed together, it can happen that one and the same combination of Family Interview Number and Sequence Number is assigned to two or even more different individuals. This will occur, for example, if different individuals each of them being Head of a different family hand in their questionnaire as 5th person in different years. In this case all the Heads would have an Interview Number of 5 and a Sequence Number of 1. A unique identification is thus impossible.

The lack of unique identification makes it impossible to analyze the determinants of risk-attitudes for individuals as well as families.

Solution

Identification variables are added, one to identify families and one to identify individuals. Both are numbered consecutively, where first, families are sorted in ascending order by survey wave, moving year, Family Interview Number. Second, individuals are additional sorted by Sequence Number.[150] This results in 833 individuals forming 320 families that can be surveyed.[151]

[149] The Sequence Number ranks people according to their position within the family hierarchy (see Institute for Social Research, Survey Research Center, University of Michigan (2011d), Paragraph C).
[150] The variable of unique identification of migrant families is named „STD_FamilyMoverID“, the corresponding variable for individual migrants is named „STD_PersonalMoverID“.
[151] The SPSS data file including all identification numbers can be found under „\Mover\(4) Personal and Family Mover IDs\4f_New Family and PersonalIDs.sav”.

3 Data cleaning of the American Community Survey

3.1 Overview of data cleaning problems

The objective of the second step of data analysis is to estimate income parameters for each year in each possible destination state depending on socio-economic characteristics gender, age and education. This is associated with the following problems outlined below.

General Problem 1: Quantification of income (see Part B, Section 3.2)

So far the terms "income" and "income parameters" have been used without concreteness. This raises several questions:

1.) Whose income is relevant in the family context (see Part B, Section 3.2.1)?
2.) What type of income is relevant (see Part B, Section 3.2.2)?
3.) Should income possibilities be estimated from all residents or all people working in that state (see Part B, Section 3.2.3)?

General Problem 2: Making ACS variables comparable to the PSID coding (see Part B, Section 3.3)

In order to model the migration decision income parameters estimated from the American Community Survey must be merged with corresponding migrants identified by the Panel Study of Income Dynamics. Because content and file structure of the two data sources are totally different from one another, key variables must be made comparable, i.e., must be coded identically.

General Problem 3: Reducing socio-economic groups (see Part B, Sections 3.4, 3.5, and 3.6)

To model the migration decision income parameters depending on socio-economic characteristics gender, age and education must be estimated for every U.S. state over 10 years from 2000 to 2009. This results in over 590,000 combinations to be estimated all of which have to have at least a sample size of 30 test persons. This requirement cannot be met even by the high density sample ACS. In order to be able to perform a valid estimation of income parameters, socio-economic groups must be reduced.

The following sections will discuss solutions to these problems in detail.

3.2 Quantification of income in the ACS

This section is concerned with general Problem 1 outlined in Part B, Section 3.1.

3.2.1 Family income by socio-economic characteristics

Problem 1: Sample size of families with same socio-economic characteristics too small

To estimate family income of a certain family in all possible destination states, a sample of at least 30 families with exactly the same combinations of gender, age and education must exist in every U.S. state. Unfortunately, even the high density sample American Community Survey does not provide sample sizes being large enough to estimate the income parameters of whole families with these very characteristics for every state and year. Thus, family income as reported by the American Community Survey cannot be applied here.

Solution 1

Family income can also be estimated by aggregation over each family member's individual income. This means, American Community Survey income variables on the individual level are applied here rather than family or household level income variables.

Problem 2: Whose income is relevant for the migration decision?

The theoretical migration decision model assumes that Head is taking the migration decision based on family income. This raises the question of who contributes to family income. Certainly Head as main earner is relevant but who else should be considered? For example, income possibilities of minor children which only earn some extra pocket money after school certainly do not trigger migration. This consideration raises the question of who is actually part of the maximization calculus in a family with more than one person.

Solution 2[152]

In my analysis the number of people adding to the family income is restricted to the two main earners. This means, Head's income is always considered. Additionally, the person having the next highest Sequence Number after the move is also considered, if he is at least 16 years old, was

[152] The algorithm can be found under „\Alpha\(2) Family Alpha aus Maple\(3) Datenmasken relevante Familieneinkommmen Head, Wife\Berechnungszubehör\Makro_Datenmaske Familieneinkommen Head, Wife.xslm."

respondent to the interview after the move and is actually living in the family. If this is not the case, the family member with the next highest Sequence Number is checked for the same criteria.

The reasons for this selection are as follows: First, children being younger than 16 are eliminated for reasons outlined above. Second, inherent to non-respondents is that data about them is not available in the data set. Thus, their income possibilities cannot be estimated. Third, income possibilities are considered only of those living with the family because people who left the family or died probably do not add to family income anymore. Often family members do not live with the family but in an institution like facilities of the armed forces, health care facilities (e.g.,hospitals), education facilities (e.g., college dormitories) or prisons.[153] Even though family members in an institution may earn money which adds to the family income, their state of residence is where the institution is located. If the rest of the family moves, but not the person living in an institution, income possibilities of the latter do not change. Consequently, income possibilities of family members living in an institution cannot trigger migration of the rest of the family. It is therefore not reasonable to include their income possibilities when the Head decides on economic migration based on family income.

3.2.2 Earned, self-employed, total personal, or welfare income

<u>Problem</u>

On the individual level the American Community Survey offers different types of income which can be divided into earned income, welfare income, retirement income and income from capital investments.[154] Furthermore, an aggregated measure of income is available, that reports "total pre-tax personal income or losses from all sources for the previous year"[155].

[153] See Institute for Social Research, Survey Research Center, University of Michigan (2010c), p. 3.
[154] For an overview of all income categories see Minnesota Population Center, University of Minnesota [No date c], online dictionary on variables concerning "Personal Income" of which "INCEARN", INCWELFR", "INCRETIR" and "INCINVEST" are mentioned here explicitly.
[155] Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable "INCTOT", paragraph on "Description".

Solution

For my analysis “total pre-tax personal income or losses from all sources for the previous year“[156] is used because it sums up all types of income. The rationale behind this is that is not clear what type of income may trigger migration.

3.2.3 Income by residents or employed persons in a U.S. state

Problem

In general, the income opportunities somebody faces in a certain U.S. state can either be estimated taking total personal income of all residents in the respective state or it can be estimated by considering only those working in that state.

Solution

There are several arguments for taking residents into account rather than people being employed in that state.

(1) Congruent universe

The universe from which income parameters are estimated must have the same characteristics as those of the migrants. Migration in my analysis is defined as changing residence rather than place of work. Consequently, state of residence must be the critical characteristic.

(2) All residents, not only employed people

Estimating income parameters based only on those individuals employed in a particular state would imply a focus on income earned, neglecting unemployment. Even if this measure would be supplemented by taking into account the welfare payments of unemployed in that state, the estimation would still be biased as employed and unemployed people do not have a congruent state of residence. An example for generating a great bias this way, is a large area state, which has only one economically strong city at the border to another state, while the rest of the state has high unemployment rates. If the majority of people working in this border town come from the neighbor state, their income creates a bias towards a higher mean income than the actual population in that large area state has. Vice versa the income parameters of the neighbor state will have a downward bias if only those actually working in the neighbor state are considered.

---

[156] Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable “INCTOT”, paragraph “Description”.

(3) Commuting activity included

Estimating the income parameters of residents assures that the commuting activity of people living in that state is also taken into account since a potential migrant can also commute to neighbor states. This argument may become clear if one thinks about commuter belts around big cities. Often highly paid employees live in commuter belts where living conditions are better increasing average income there. In contrast, those people living in the city itself may not earn so much money on average. Therefore it could be that average income in the commuting belt is higher than in the city itself. If a forecast about income for residents in the city is needed, it would be wrong to take all people working there into account rather than only looking on actual residents.

Therefore, income parameters are estimated based on all residents of a particular state.

3.3 Making ACS variables comparable to the PSID coding[157]

This section is concerned with general Problem 2 outlined in Part B, Section 3.1.

3.3.1 Problem and solution

Problem

Income parameters depend on the socio-economic characteristics of the specific migrant. This means, migrants and income parameters are merged by variables on socio-economic characteristics. The merging can only be performed if and only if variables on age, gender, education and U.S. states have identical coding schemes in both datasets. This is not the case (see Table 15, p. 99, Table 16, p. 99 and Table 17, p. 100).

Solution

Raw data files of the American Community Survey have to be recoded according to the standardized coding scheme used in the Panel Study of Income Dynamics. While **age** is coded in the same way in both datasets, variables on **gender and U.S. state** only have the same answer options but different codes. Thus, parameter values simply have to be recoded. For example, Virginia is coded (51) in the

[157] The resulting standardized American Community Survey files of each wave are named e.g., “1_stdState_ACS_2000.sav” for wave 2000 and are located in folder „\American Community Survey\(1) Transform State Code FIPS to PSID”. The SPSS-Syntax file for the standardization procedure including further comments is named „(1) Transform State Code FIPS to PSID.sps“ and can be found in the same folder.

ACS files but must be recoded to (45) to match the standardized PSID variable coding scheme. Gender is coded 0,1 versus 1,2.

Concerning **education** standardized PSID variables and ACS variables do not have the same parameter values. Education categories “no education” and “Doctorate or Ph.D.” which exist in the Panel Study of Income Dynamics are not separately defined in the American Community Survey. This means for both education levels no corresponding income parameters can be estimated. Instead migrants with “no education” must be merged with the next higher category “8th grade or less” and individuals with a Doctorate or a Ph.D. must be merged with those having a Master or Professional degree since this is the highest education level defined in the ACS (see Table 17, p. 100).

3.3.2 SPSS-implementation

U.S. State Codes

The American Community Survey offers two coding schemes for U.S. states: the coding scheme of the Inter-University Consortium for Political and Social Research (ICPSR) and the coding scheme of the Federal Information Processing Standards (FIPS).[158] Both are not compatible to the PSID-coding that was used in the Panels Study of Income Dynamics. Because the FIPS-coding is more similar to the PSID-coding it is recoded according to the PSID coding scheme.

Before recording can be performed, it must be noted that the various documentations available for FIPS and PSID coding contradict each other for two states (bold typed in Table 15, p. 99):

- **PSID-coding of Oregon:** In the PSID documentation of 1981 Oregon is coded with 36[159], while the PSID documentation of 1985 states a code of 41[160]. The latter source must be wrong because code 41 stands for the state of Tennessee in both sources.[161] Therefore, Oregon is PSID-coded by 36.

- **FIPS-coding of Vermont:** The documentation of FIPS state codes can be found in the PSID documentations. In the PSID documentation of 1985 Vermont is coded with 59.[162] This must be wrong for two reasons. First, PSID codebooks of Single-Year Family Files of all waves from 1989

---

[158] The corresponding variables are named STATEICP and STATEFIP in all waves.
[159] See Institute for Social Research, Survey Research Center, University of Michigan (1982), p. 629.
[160] See Institute for Social Research, Survey Research Center, University of Michigan (1988), p. 714.
[161] See Institute for Social Research, Survey Research Center, University of Michigan (1982), p. 630 and Institute for Social Research, Survey Research Center, University of Michigan (1988), p. 715.
[162] See Institute for Social Research, Survey Research Center, University of Michigan (1988), p. 717.

to 2009 report that FIPS codes only range from 1 to 56, making a code of 59 impossible.[163] Second, codes are reported in ascending order with code 49 mentioned directly before and code 51 mentioned directly after code 59.[164] Therefore it seems plausible that Vermont is FIPS-coded 50. Finally, ACS raw data files reveal that Vermont is actually FIPS-coded 50.

After shedding light on the contradicting documentations for Oregon and Vermont the recoding of the FIPS state codes according to the PSID state codes is performed as follows:

PSID-Code	State	FIPS-Code	State
1	Alabama	1	Alabama
2	Arizona	2	Alaska
3	Arkansas	4	Arizona
4	California	5	Arkansas
5	Colorado	6	California
6	Connecticut	8	Colorado
7	Delaware	9	Connecticut
8	District of Columbia	10	Delaware
9	Florida	11	District of Columbia
10	Georgia	12	Florida
11	Idaho	13	Georgia
12	Illinois	15	Hawaii
13	Indiana	16	Idaho
14	Iowa	17	Illinois
15	Kansas	18	Indiana
16	Kentucky	19	Iowa
17	Louisiana	20	Kansas
18	Maine	21	Kentucky
19	Maryland	22	Louisiana
20	Massachusetts	23	Maine
21	Michigan	24	Maryland
22	Minnesota	25	Massachusetts
23	Mississippi	26	Michigan
24	Missouri	27	Minnesota
25	Montana	28	Mississippi
26	Nebraska	29	Missouri
27	Nevada	30	Montana
28	New Hampshire	31	Nebraska
29	New Jersey	32	Nevada
30	New Mexico	33	New Hampshire
31	New York	34	New Jersey
32	North Carolina	35	New Mexico

[163] See for example, Institute for Social Research, Survey Research Center, University of Michigan (2011c), p. 2.
[164] Institute for Social Research, Survey Research Center, University of Michigan (1988), p. 717.

33	North Dakota	36	New York
34	Ohio	37	North Carolina
35	Oklahoma	38	North Dakota
36	**Oregon**	39	Ohio
37	Pennsylvania	40	Oklahoma
38	Rhode Island	41	Oregon
39	South Carolina	42	Pennsylvania
40	South Dakota	44	Rhode Island
41	Tennessee	45	South Carolina
42	Texas	46	South Dakota
43	Utah	47	Tennessee
44	Vermont	48	Texas
45	Virginia	49	Utah
46	Washington	**50**	**Vermont**
47	West Virginia	51	Virginia
48	Wisconsin	53	Washington
49	Wyoming	54	West Virginia
50	Alaska	55	Wisconsin
51	Hawaii	56	Wyoming
0	Other territories	0	Other territories

Table 15: Recoding of FIPS-coding in order to accord with PSID-coding of U.S. states.
Source: See for the PSID-coding Institute for Social Research, Survey Research Center, University of Michigan (1982), p. 611-636; See for FIPS coding Institute for Social Research, Survey Research Center, University of Michigan (1988), p. 701-719 and ACS raw data files waves 2000 to 2009.

<u>Gender</u>

In the American Community Survey gender is coded with (1) and (2) instead of (0) and (1) in the standardized variables used in the Panel Study of Income Dynamics. Therefore, the variable in the American Community Survey is recoded according to the PSID-coding as follows:

Original ACS variable SEX		Standardized PSID variable STD_SEX	
Code	**Label**	**Code**	**Label**
1	Male	0	Male
2	Female	1	Female
		9999	Missing

Table 16: Recoding of ACS coding of gender to accord with PSID coding.
Source: For coding of the ACS variable Sex see Minnesota Population Center, University of Minnesota [No date c].

<u>Education</u>

The American Community Survey education variable “indicates respondents' educational attainment, as measured by the highest year of school or degree completed”[165] of all respondents of the

[165] Minnesota Population Center, University of Minnesota [No date c] on variable EDUC, sheet “Description”.

American Community Survey.[166] The coding scheme is different from the standardized variables on education that was created for the Panel Study of Income Dynamics. Therefore, certain harmonization must be undertaken. Table 17, p. 100 compares the coding scheme of the ACS variable on education (left column of Table 17, p. 100) to the coding scheme of the standardized education variables used in the PSID data (right column of Table 17, p. 100). Corresponding answers of the ACS variable and standardized PSID variable are reported in the same row with some categories not having a counterpart.

<table>
<tr><th colspan="2">Original ACS variable on education</th><th colspan="2">Standardized PSID variable on education</th></tr>
<tr><th>Code</th><th>Label</th><th>Code</th><th>Label</th></tr>
<tr><td></td><td>---</td><td>0</td><td>No education</td></tr>
<tr><td>0</td><td>N/A or no schooling</td><td>9999</td><td>Missing</td></tr>
<tr><td>1</td><td>Nursery school to grade 4</td><td rowspan="2">1</td><td rowspan="2">8th grade or less</td></tr>
<tr><td>2</td><td>Grade 5, 6, 7, or 8</td></tr>
<tr><td>3</td><td>Grade 9</td><td rowspan="3">2</td><td rowspan="3">9th to 11th grade, no graduate</td></tr>
<tr><td>4</td><td>Grade 10</td></tr>
<tr><td>5</td><td>Grade 11</td></tr>
<tr><td>6</td><td>Grade 12</td><td rowspan="2">3</td><td rowspan="2">High school graduate</td></tr>
<tr><td>7</td><td>1 year of college</td></tr>
<tr><td>8</td><td>2 years of college</td><td>4</td><td>Associate's degree</td></tr>
<tr><td>9</td><td>3 years of college</td><td colspan="2">---</td></tr>
<tr><td>10</td><td>4 years of college</td><td>5</td><td>Bachelor’s degree</td></tr>
<tr><td>11</td><td>5+ years of college</td><td>6</td><td>Master/Professional Degree</td></tr>
<tr><td colspan="2">---</td><td>7</td><td>Doctorate, Ph.D.</td></tr>
</table>

Table 17: Recoding of ACS education variables to accord with PSID coding with corresponding answers reported in the same row.
Source: For coding of the ACS variable EDUC see Minnesota Population Center, University of Minnesota [No date c].
“---“ denotes no corresponding answer existent; “N/A” denotes not available. Note, that the category “3 years of college” is not available in ACS waves 2000 to 2009.

First, the **lowest education level in the PSID data (“no education”)** does not exist separately in the ACS variable on education. Instead, it is aggregated with missing data (“N/A or no schooling“, where N/A denotes not available). This mixture of answers makes it impossible to differentiate between individuals with no education and those with missing data. Consequently, income parameters of those having no education cannot be estimated. The ACS answer “N/A or no schooling“ cannot be applied to any education category defined by the standardized education variables and is therefore recoded as missing information.

Second, the next higher education levels, namely **up to 12th grade** can easily be transformed from the ACS definition to the standardized variable definition as can be seen in Table 17, p. 100 where corresponding answers of the ACS and standardized variable are reported in the same row: ACS-

[166] See Minnesota Population Center, University of Minnesota [No date c] on variable EDUC, sheet “Universe”.

Categories “(1) Nursery school to grade 4” and “(2) Grade 5, 6, 7, or 8” are recoded to “**(1) 8th grade or less**”. ACS-Categories (3) to (5) are recoded to “**(2) 9th to 11th grade, no graduate**”. Finally ACS-categories “(6) Grade 12” and “(7) 1 year of college” are recoded to “**(3) high school graduate**”.

Third, for college education the ACS variable description reveals that it assigns “each degree the number of years it typically takes: ‘2 years of college’ for an associate's degree; ‘4 years of college’ for a bachelor's degree; and ‘5+ years of college’ for a graduate or professional degree”.[167] Obviously one year of college is not considered to have any equivalent college degree as nothing is said about it in the variable description. Therefore, ‘1 year of college’ is recoded as high school graduate. This fits the view explained in Part B, Section 2.3.3.3.2.3 before when PSID data was standardized. Correspondents to associate’s, bachelor’s and master’s degrees are exactly defined by the ACS variable description and are recoded that way.

Fourth, the highest level of the standardized education variable “**Ph.D./Doctorate”** has no corresponding answer in the ACS variable but seems to be included in the category “5+ years of college” implicitly. Therefore, separate income parameters of individuals with that degree cannot be estimated from the American Community Survey.

Education gained outside the U.S. had to be reported as U.S. equivalents in the American Community survey. Thus the ACS variable includes foreign as well as U.S. education.

Total Personal Income

The ACS variable INCTOT “reports each respondent's total pre-tax personal income or losses from all sources for the previous year“.[168] The only recoding necessary here is the coding for missing values from (9999999) in the ACS files to (9999) like in the PSID files.

3.4 Reducing socio-economic groups

This section is concerned with general Problem 3 outlined in Part B, Section 3.1.

Problem

The number of socio-economic groups for which income parameters have to be estimated in order to model the migration decision adds up to 587,520 from 2000 to 2009: With gender having two categories (male, female), age having 96 levels (ages 0 to 95 are reported in the ACS) and education

[167] Minnesota Population Center, University of Minnesota [No date c] on variable EDUC, sheet “Comparability”.
[168] Minnesota Population Center, University of Minnesota [No date c] on variable INCTOT, sheet “Description”.

having 6 levels in the ACS files there are all together 1,152 socio-economic groups in each of the 51 U.S. states over 10 years. Even though the American Community Survey is a high density sample, the sample size in each group is not high enough to estimate income parameters, sometimes not having a single respondent in the survey (e.g., male person in Idaho being 25 years old and having a master's degree or equivalent in 2000). Without a reliable estimation of income parameters the migration decision cannot be empirically analyzed. In order to achieve the sample size needed for a robust estimation of income parameters, the originally 587,520 socio-economic groups must be reduced by clustering people that are similar concerning their income parameters. This raises two problems: First, along which dimensions can the grouping be performed (years, U.S. states, gender, education and age) and second, which method can be used.

Solution

To get an economic intuition on which dimensions can be grouped together (different years, different U.S. states, men and women, different education levels, and different ages) a descriptive analysis on mean and variance of income is run. Insights gained are then used to apply a cluster analysis. Finally, economic considerations are used to further group people.

3.4.1 Variables to be clustered

Of all variables over which a clustering could be performed, U.S. states and years can be ruled out:

In the theoretical model migration is triggered by differences in income between U.S. states. Therefore a clustering along U.S. states, that means clustering the sample of socio-economic identical individuals (same gender, age, education) of ,for example, Washington D.C. and Idaho, and applying the same income parameters to both states, does not make sense since it eliminates income differences between states.

Clustering over year seems also not reasonable because of structural differences over the 10 years surveyed. This leaves the socio-economic characteristics gender, age, and education as potential candidates for clustering.

3.4.2 Simultaneous versus separate clustering of variables

To reduce the remaining socio-economic groups, I perform a ceteris paribus clustering **for each variable separately** (gender, education, and age). A simultaneous clustering of people who are similar concerning **all variables** (gender, education, and age) is not reasonable because it would mean a clustering over all groups: For example, the same income parameters are applied to a 15 year old boy in New York in 2007 and a 92 year old woman in Alaska in 2009. This example already shows the first drawback of such a procedure. It could – and probably would because of the enormous number of groups – result in a clustering that makes no sense from an economic perspective. Neither are the socio-economic characteristics of people in the same cluster comparable nor their reasons to move. A second drawback exists: Clustering over all around 600,000 groups is statistically difficult. In principle, this could be done by several pair wise mean-difference tests.[169] The problem of multiple two-sample tests is the accumulation of probability of error, which becomes already unacceptably high for relatively small numbers of groups.[170] With 587,520 groups 172,589,581,440 pair wise tests would be necessary which will add up to almost 100% probability of error if the significance level for a single test would be only 1%.[171] Obviously, the level of significance is dramatically rising with the number of pair wise tests and is therefore no option in this study with almost 600,000 groups. Because of the high number of socio-economic groups and for economic reasons the only aggregation which makes sense here is a ceteris paribus clustering separate for each variable.

3.4.3 Descriptive analysis of income by gender, education, and age[172]

The standard input parameters of the migration decision model are mean and variance of income. This means, that only groups that have similar means and variances of income can be clustered. To get an idea how gender, education and age can be clustered, mean and variance of income for each of these characteristics is analyzed over all other variables.

To understand the graphs discussed in the next section, characteristics of variables on age and income must be elaborated: First, note, that age is top coded “in order to protect the confidentiality of respondents“[173]. This means, individuals at the extreme upper end of 90 years or older will not be reported with their true age but with the state median of all values exceeding the top codes.[174] Depending on the U.S. state, the median lies between 92 and 94 years. Second, income is only

[169] A multiple mean-difference test tests the null hypothesis $H_0: \mu_1 = \mu_2 = \dots = \mu_n$ which can only reveal that at least one of the samples has a mean different from the others. But it does not give an answer to the question which sample is similar to the other and which one is different (see Rasch, Friese, Hofman, and Naumann (2006), p. 27).
[170] The cumulative probability of error adds up to $\alpha_{cum} = 1 - (1 - \alpha_{test})^m$ with α_{test} being the probability of error for each single pair wise test and m being the number of pair wise tests (see Rasch, Friese, Hofman, and Naumann (2006), p. 3.)
[171] The number of necessary pair wise tests m is a function of groups to be compared k and can be calculated by $= \frac{k \cdot (k-1)}{2}$ (see Rasch, Friese, Hofman, and Naumann (2006), p. 4).
[172] All graphs and source codes in SPSS can be found in folder „\American Community Survey\(2) Descriptive Analysis of Gender, Age and Education”.
[173] Minnesota Population Center, University of Minnesota [No date e].
[174] For more details about top codes, its meaning, and the actual top codes see Minnesota Population Center, University of Minnesota [No date e].

reported for people aged 15 and older[175] and is also top coded to 484,900 U.S. Dollar.[176] Again, if income exceeds this value, the state median of all records exceeding the top code is reported.

Finally, to get a representative sample of the U.S. population, weights as offered by the American Community Survey are applied for the descriptive analysis. All graphs presented exemplarily show data from the American Community Survey 2001. Results presented here also hold for unweighted means and variances.

3.4.3.1 Gender

The objective of the descriptive analysis is to figure out whether men and women can be clustered. This can be done if men and women are pair wise similar concerning (weighted) mean and (weighted) variance of income over all variables. Therefore, weighted mean and weighted variance of income depending on gender will first be analyzed over all ages, then over all education levels and finally over all U.S. states. If men and women are similar in (weighted) mean and (weighted) variance of income in all pair wise comparisons, then men and women can be grouped.

---

[175] See Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable "INCTOT", paragraph "Universe".

[176] Total personal income gives the sum from all types if income that are each top coded. These are INCWAGE (wage and salary income), INCBUS (non-farm business income), INCSS (Social Security income), INCWELFR (welfare, public assistance income), INCSUPP (Supplementary Security Income), INCINVST (interest, dividend, and rental income), INCRETIR (retirement income), INCOTHER (other income) (see Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable "INCTOT", paragraph "Comparability").

Income by gender over age

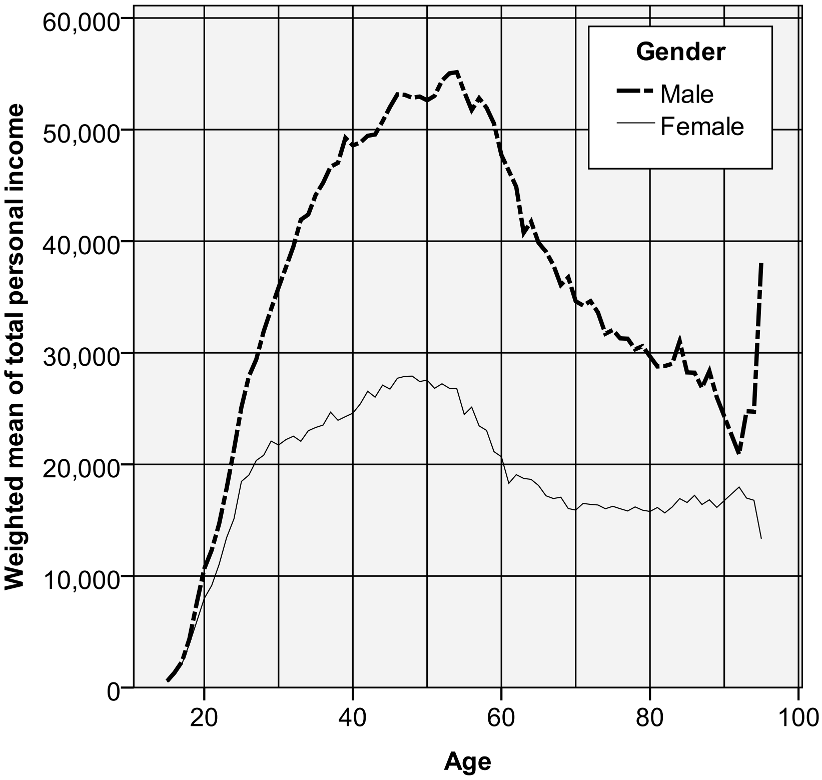


Figure 7: Weighted mean of income by age for men and women in 2001. Cases are weighted according to the ACS personal weights in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

Figure 7, p. 105 gives the weighted mean income of men and women aged 15 or older exemplary for ACS wave 2001.[177] While weighted mean income is almost the same for boys and girls aged 15, the difference gets greater each year until people turn around 50 years old with men earning considerably more on average than women. Even though this difference gets a little bit smaller after age 50, it remains enormous during whole working ages (age 16 to the statutory retirement age of 65[178]) and even further. This phenomenon can be seen in all ACS waves. Only the weighted mean income for ages 85 differs slightly from wave to wave, while men always keep their lead except for the year 2000 where curves cross each other for certain ages higher than 85. The reason for the different structure of weighted income for people aged 85 and older is due to the declining samples size for older people and the top coding of age in the American Community Survey.

[177] The ACS reports income only for people aged 15 or older (see Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable “INCTOT”, paragraph “Universe”). The oldest person in the the ACS sample is aged 95 (see Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable “AGE”, paragraph “Codes”).

[178] Depending on year of birth the full retirement age is 65 or 66 (see U.S. Social Security Administration (2011)).

The retaining conclusion is that men have a considerably higher weighted mean income than women, especially during their working ages. The weighted variance of income of men and woman also differs systematically as can be seen in Figure 8, p. 106.

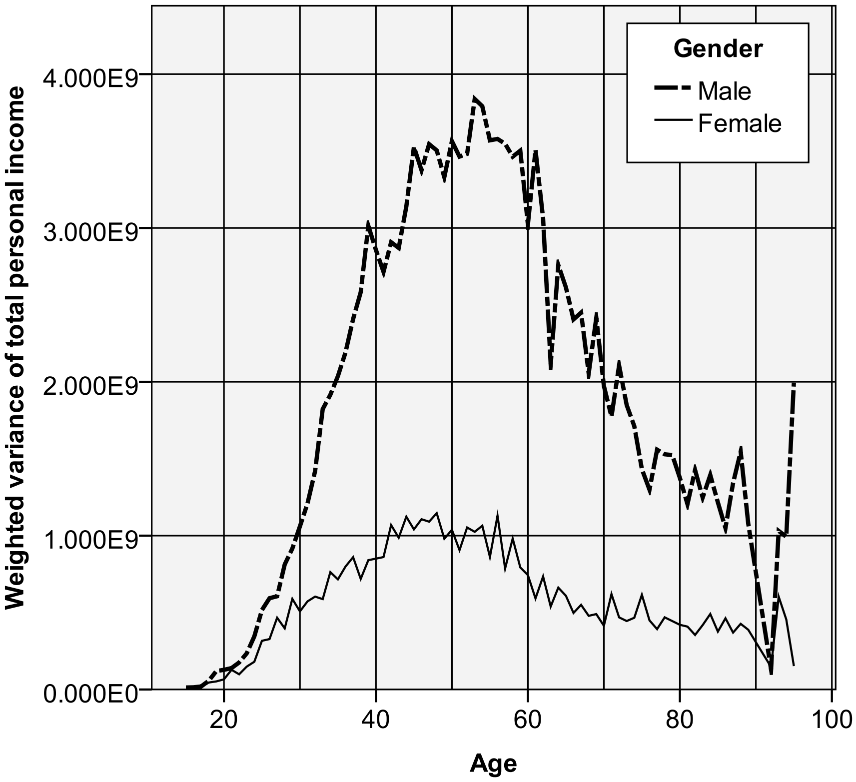


Figure 8: Weighted variance of income by age for men and women in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

Men having higher weighted mean income also have a higher weighted variance of income than women for all ages until they turn around 85. This can be seen for all waves from 2000 to 2009. Again for people aged 85 and older graphs look different from wave to wave which is due to the small sample size in these ages and the top coding.

<u>Income by gender over education levels</u>

Turning to education, weighted mean and weighted variance of income for a given education level are always greater for men than for women while this difference gets greater with every education level after 9^{th} to 11^{th} grade. This applies to all waves of data with no exception. Figure 9, p. 107 and Figure 10, p. 107 show the respective graphs exemplary for ACS wave 2001.

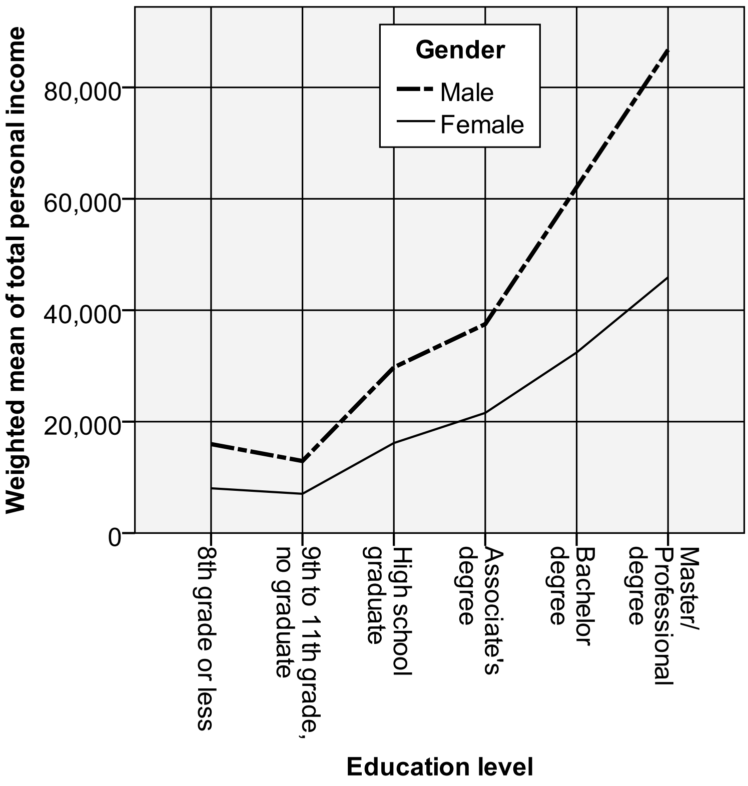


Figure 9: Weighted mean of income by highest education level achieved for men and women in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

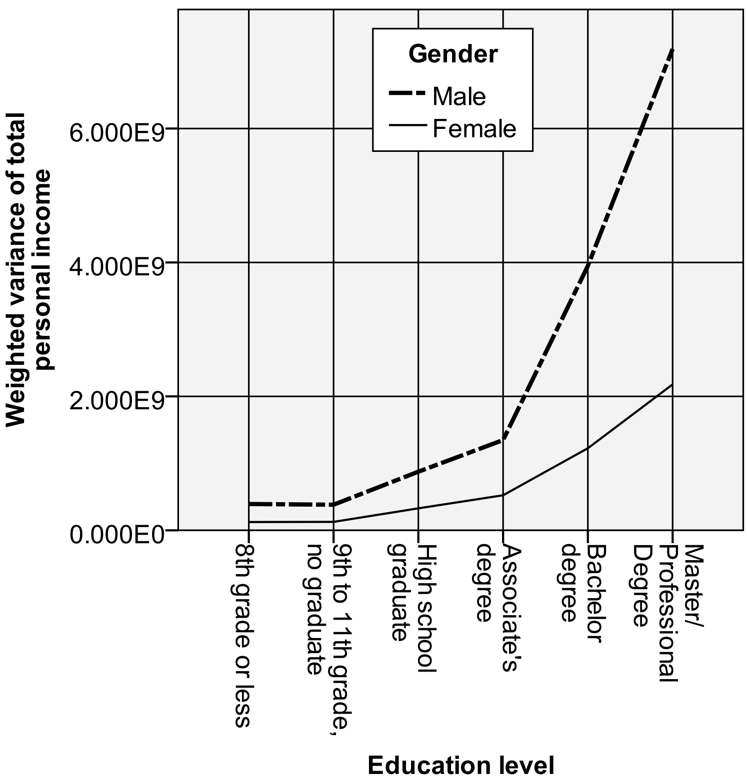


Figure 10: Weighted variance of income by highest education level achieved for men and women in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

Income by gender over state of residence

The huge difference in weighted mean and weighted variance of income between men and women can also be observed over all U.S. states in all ACS waves from 2000 to 2009 with no exception. Men earn higher wages on average but also have a higher weighted variance of income. Figure 11, p. 109 shows the respective graphs exemplary for ACS wave 2001.

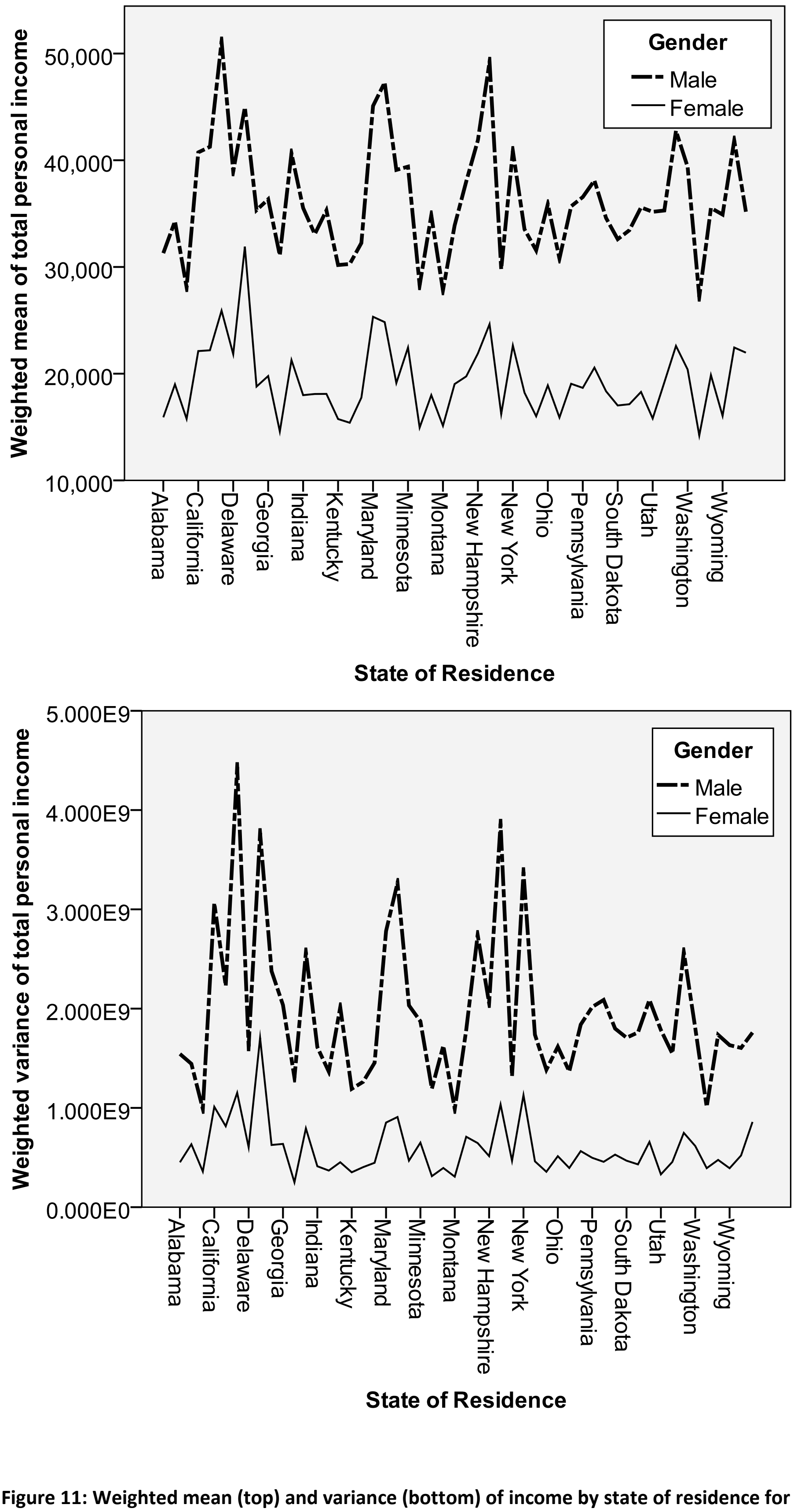


Figure 11: Weighted mean (top) and variance (bottom) of income by state of residence for men and women in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population. Source: Own calculations based on data of the American Community Survey 2001.

In summary, the descriptive analysis of income by gender clearly shows that men have a considerably higher weighted mean and weighted variance of income than women for all ages[179], education levels, and U.S. states over all years. The difference between men and women gets larger the higher the education level is and is in particular pronounced between age 15 and 50. This shows that income parameters for male and female individuals must be estimated separately rather than being clustered to one sample at least for their working ages. A grouping of socio-economic groups along gender does not seem plausible.

3.4.3.2 Education

Income by education over age

The graphical analysis of income by education over all ages (see Figure 12, p. 111) must be interpreted with some caution for people up to 20. The reason is that higher academic degrees are associated with a certain minimum age as it takes several years of academic studies to achieve them. For example, it is unrealistic that people of 18 already have a master's degree.

Figure 12, p. 111 clearly shows that weighted mean income systematically rises with education level – the lowest curve gives weighted mean income of the lowest education level (8th grade or less), each next higher curve standing for the next higher education level with the upper curve giving weighted mean income of the highest education level (Master/Professional degree).

Regarding similarities, it becomes clear that weighted mean income for the two lowest education levels (8th grade or less and 9th to 11th grade) is most similar. Furthermore, differences in weighted mean for having a high school diploma and having an associate's degree have a tendency to be relatively close compared to higher levels of education for working ages.

[179] With only one exception for weighted mean income in 2000 for certain ages over 85. This effect is due to the small sample size for people in that age.

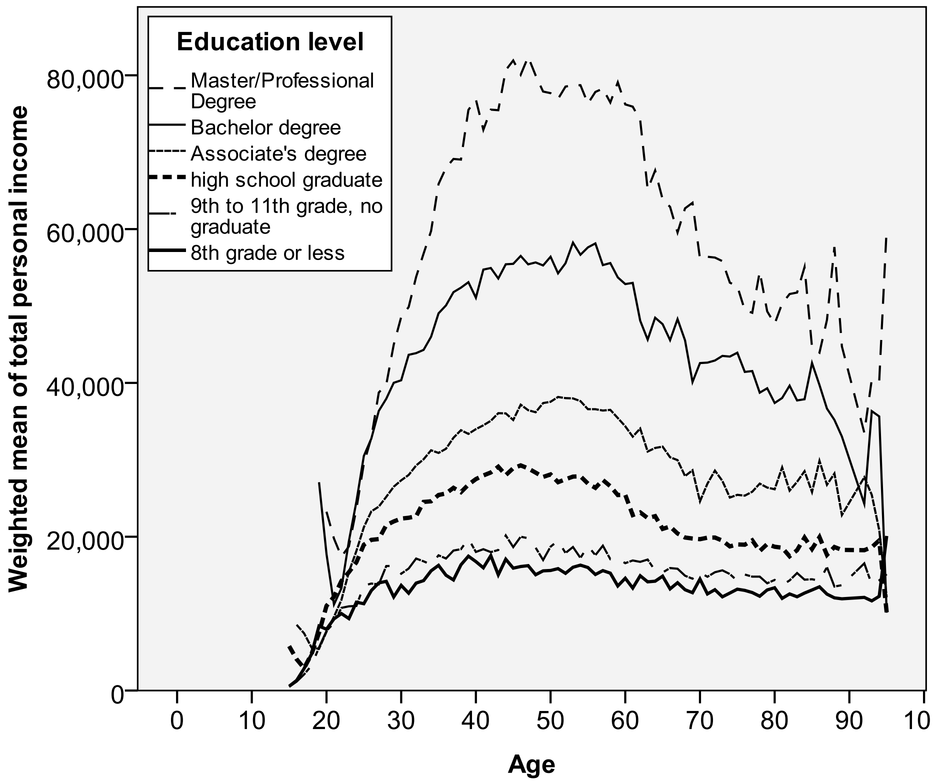


Figure 12 : Weighted mean of total personal income by education level in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

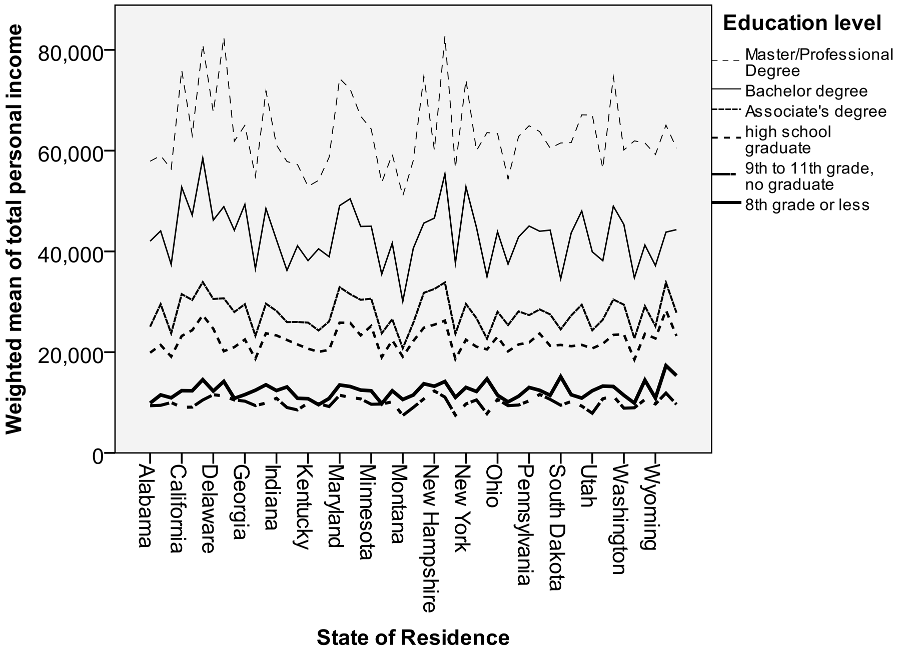
Graphs giving weighted variance of income rather than weighted mean income qualitatively look alike for all ACS waves. They give a first indication that for working ages differences in weighted mean and weighted variance of income between education levels get the higher the higher the actual level is. This relationship does not hold for people under 20 and those over 80. While younger people systematically cannot achieve all education levels, the unsystematic curves for older people could be due to cohort effects and/or small sample sizes and top coding.

<u>Income by education over gender</u>

The relationship between education and gender was already analyzed above in Figure 9, p. 107 and Figure 10, p. 107. It can be seen that the results so far also hold for men and women separately. That is, weighted mean and weighted variance of income are the higher the higher the education level of a person is (true for both men and women), while the difference between education levels is smallest for the lowest education levels (between having finished 8th grade or less and 9th up to 11th grade).

Income by education over state of residence

Finally, education income is analyzed for all U.S. states. Figure 13, p. 112 illustrates that the lower the education levels, the closer weighted mean and weighted variance of income are to the next higher level of education. This effect is even more pronounced for weighted variances.



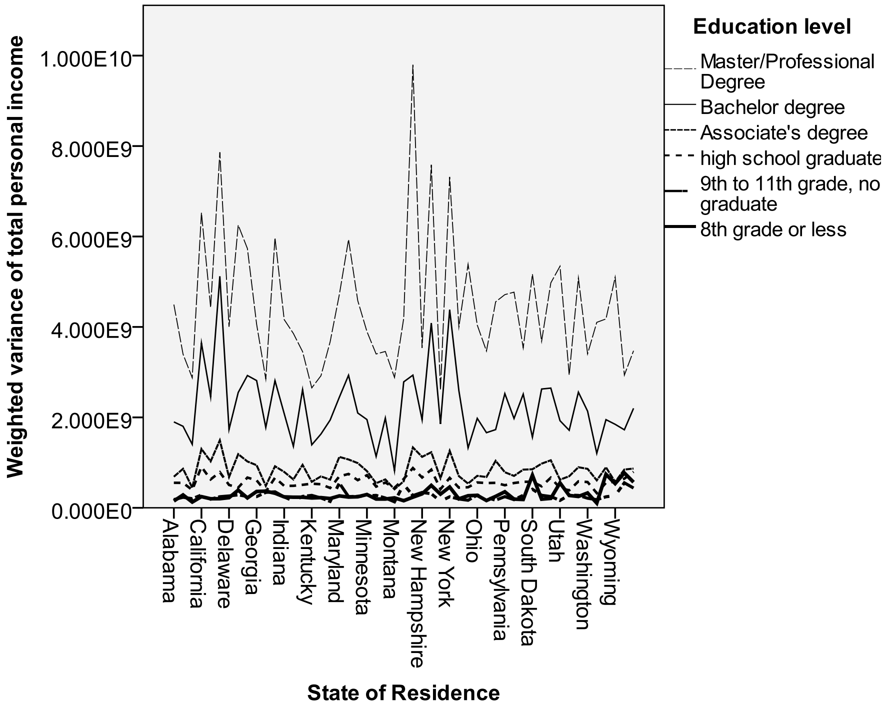


Figure 13: Weighted mean (top) and variance (bottom) of total personal income by education level for men and women in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

It can be concluded that grouping along education levels can at best be performed for lower levels of education.

3.4.3.3 Age

Grouping of people along age seems reasonable from an economic perspective. This view is supported by the graphical analysis of income over age by education (see Figure 12, p. 111) and by gender (see Figure 7, p. 105 and Figure 8, p. 106) discussed above. Another illustration of income by age is given in Figure 14, p. 113. The logarithm of weighted mean income shows that the closer age groups are to each other, the smaller the difference in weighted mean income is. Where individuals have a higher percentage rise in weighted mean income year by year for ages 15 to around 45, with relatively stable percentage growth in weighted mean income in the middle of their work-life, and falling weighted mean income afterwards (see also Figure 7, p. 105).

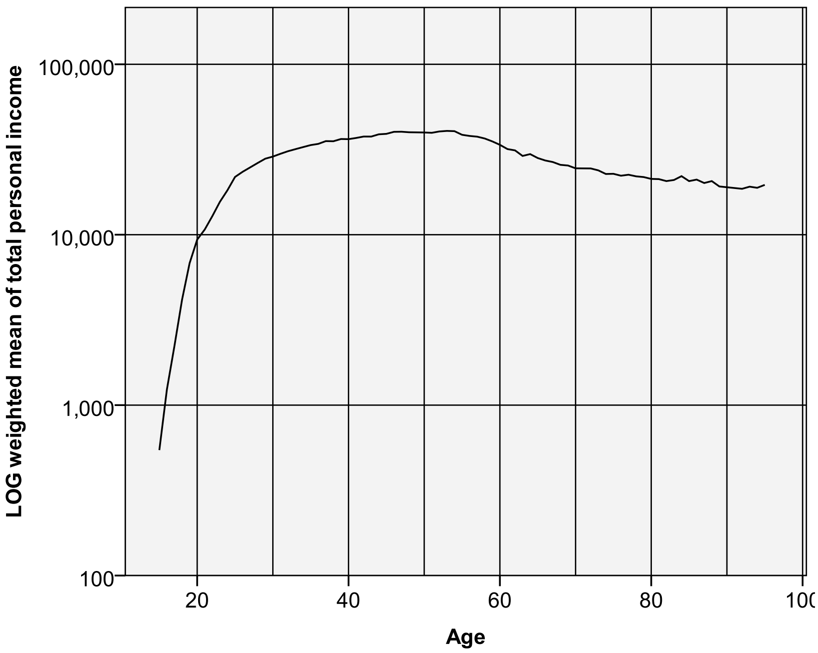


Figure 14: Log weighted mean of total personal income by age in 2001. Cases are weighted according to the ACS personal weight in order to get a representative sample of the U.S. population.
Source: Own calculations based on data of the American Community Survey 2001.

3.4.3.4 Conclusion of descriptive analysis

The descriptive analysis has shown that weighted mean and weighted variance of income by gender are considerably different in all age groups - especially during working ages- , U.S. states, and for all education levels. Therefore, a grouping along gender (i.e., pooling the sample of male and female individuals) does not make sense. Furthermore, it turned out that lower levels of education are associated with relatively similar weighted means and weighted variances. The same can be concluded for age groups next to each other which are also very similar concerning the relevant income parameters. Consequently, the grouping of socio-economic groups is performed over the dimensions age and education.

3.4.4 Cluster analysis

To reduce the number of socio-economic groups by aggregation over age and education separately, a cluster algorithm is applied. Its goal is to group objects into clusters that are similar in terms of certain variables, while the difference between clusters should be as great as possible.[180] This is exactly what must be done to reduce socio-economic groups and achieving a larger sample size in each group. Note that all explanations of this section equally refer to weighted and unweighted mean and variances.

3.4.4.1 Clustering methods

The cluster analysis can be performed by many different methods. I chose an agglomerative hierarchical cluster analysis with the Ward-Method using squared Euclidean distance measures on z-transformed variables mean income, variance of income and age. The decision to choose this methods consists, first, of the type of proximity measure and, second, of the cluster model applied. Both are discussed in detail in this section.

In principle, two types of **proximity measures** exist, namely similarity and distance measures.[181] Because absolute difference in mean and variance of income is important, distance measures are appropriate here.[182] The right type of distance measure depends on the scale of variables considered: mean and variance of income are both scaled metrically. Of all distance measures for variables scaled metrically the squared Euclidean distance is applied because it gives greater differences of mean and variance of income a higher weight.[183]

The **cluster model** chosen is the agglomerative hierarchical cluster analysis. The reasons for choosing a hierarchical over a partitioning method are threefold:[184] First, partitioning methods need a predefined clustering which is then reorganized step by step until a predefined criterion is met.[185] Second, the algorithm often get stuck in local rather than a global optimum which makes it necessary to try different starting clusters. Third, the solution depends on the initial starting cluster chosen. In contrast to partitioning methods hierarchical methods either start with a single cluster including all records which is then split up into smaller clusters (divisive hierarchical cluster analysis) or they start

---

[180] See Bortz and Schuster (2010), p. 453.
[181] See Everitt, Landau, Leese, and Stahl (2011), pp. 2-20.
[182] See Backhaus, Erichson, Plinke, and Weiber (2008), p. 408.
[183] See Backhaus, Erichson, Plinke, and Weiber (2008), p. 405.
[184] For disadvantages of the partitioning cluster algorithms see Backhaus, Erichson, Plinke, and Weiber (2008), p. 413.
[185] See Backhaus, Erichson, Plinke, and Weiber (2008), p. 412.

with each object being a separate cluster and successively merging them into larger clusters (agglomerative hierarchical cluster analysis). The latter is chosen here because it follows the intuitive approach of considering each group separately and then merging similar clusters step by step. For this approach, the number of final cluster must be predetermined. Within the agglomerative hierarchical cluster analysis the **Ward-Method** is used because it neither has a tendency of creating larger clusters nor smaller clusters but rather equally large clusters.[186] This assures that mean and variance of income of each cluster can be estimated from almost the same sample size.

Before performing the agglomerative hierarchical cluster analysis with the Ward-Method using squared Euclidean distance measures the **variables of interest (mean and variance of income) are z-transformed**. This is necessary because mean and variance of income are not measured on the same level (variance is a squared parameter) which results in a bias by implicitly weighting differences in variance higher than those in mean income.[187] An example may clarify the problem: A difference of 100 units in variance of income is qualitatively not the same as the same difference in mean income. A difference in mean income of 100 U.S. dollar is more meaningful (i.e., implies a greater economic difference) than the same difference in variance of income. Therefore, mean and variance of income are z-transformed to assure a mean of zero and a standard deviation of one to each variable. Then a difference of one unit has the same weight for both variables.

3.4.4.2 Four different types of clustering: weighted samples and same clustering for all years?

Before starting the cluster algorithm, two principal questions must be addressed: First, it must be decided whether mean and variance should be estimated by a **weighted or unweighted sample**.[188] The American Community Survey offers personal weights that can be applied to get a sample representative of the U.S. population.[189] On the one hand, it seems valuable to apply weights in order to get a representative estimation of income parameters. On the other hand, weighting personal records is especially afflicted with the danger of bias when the sample size is small and weights are high.

[186] See Backhaus, Erichson, Plinke, and Weiber (2008), p. 424 f.
[187] See Backhaus, Erichson, Plinke, and Weiber (2008), p. 444.
[188] Note, that the descriptive analysis was performed for weighted samples only for illustration purposes.
[189] Personal weights are created to conform with population estimates of the Population Estimates Program (PEP) of the Census Bureau by geographic and demographic characteristics like gender, age, race (see U.S. Census Bureau (2009c), p. 11-1 ff.).

The second question to be answered is whether the cluster analysis should be **run for each year separately**, resulting in different clustering of people in each year, or the grouping should be the same in each year. Certainly, running a separate cluster analysis for each wave of data where mean and variance are estimated separately for each wave is the most precise approach. On the other hand, economic interpretation will get difficult if one and the same person would be grouped into different socio-economic groups in different years. For example, 32 year old individual is clustered with people aged 22 to 44 in 2004, while the same person is clustered with people being 25 to 32 years old in 2006. Obviously, the estimation resulting from the different clusters would be different. A solution could be the **pooling over years**, meaning that people are clustered into groups by their average mean and variance of income over all years.[190] This would result in stable cluster definitions for all years surveyed. Here again it is not clear upfront which approach is preferable. Therefore, both are considered as robustness checks in the analysis.

Because of the two problems discussed above the cluster analysis is performed in four different ways, each resulting in different socio-economic groups for all years from 2000 to 2009.

Cluster definition changes over year?	Same cluster definition over all waves (so called pooled approach)		Separate cluster definition in each year (so called separate approach)	
Sample weighted/ unweighted	weighted	unweighted	weighted	unweighted
Abbreviation used when referring to the approach	→ Pooled, weighted	→ Pooled, unweighted	→ Separate, weighted	→ Separate, unweighted

Table 18: Four different types of clustering: un/-weighted sample with pooled/separate cluster definition over years. Source: Own figure.

These four different types of cluster analysis are performed on age and education separately.

[190] This means, for each year mean and variance of income are estimated. The arithmetic mean of annually mean and variances is taken for "pooled clustering".

3.4.4.3 Implementation and results of separate clustering over age and education[191]

Cluster analysis for education

The descriptive analysis has suggested that the two lowest education levels are most similar, while differences in mean and variance of income get larger for higher education levels. Therefore, the originally 6 education levels are reduced to 5 and 4 cluster, respectively. The results of the cluster analyses using the agglomerative hierarchical cluster analysis with the Ward-Method using squared Euclidean distance measures with mean and variance of income being z-transformed (see Part B, Section 3.4.4.1 for details on this method) are the same for all four different types of clustering in all years (see Table 18, p. 116 for an overview of the four types of clustering).

Education levels as available in the American Community Survey[192]		Clustering of education levels to clusters in case of …	
		… 5 education clusters	… 4 education clusters
1	**8^{th} grade or less**	1	1
2	**9^{th} to 11^{th} grade**	1	1
3	**High school graduate**	2	2
4	**Associate's degree**	3	2
5	**Bachelor's degree**	4	3
6	**Master/Professional degree or higher**	5	4

Table 19: Results of cluster analysis over education for all four cluster approaches (weighted/unweighted and pooled/separate).
Source: Own calculation. All calculations can be found under path „\American Community Survey\(3) Cluster Analysis Education".

Obviously, the cluster analysis confirms the intuition of the descriptive analysis. Whether aggregation of education to five or even only four clusters is necessary to reach the sample size needed in each socio-economic group can only be decided later on when socio-economic groups are defined based on their clustered education level and age level. Therefore, the next step is aggregation over age.

[191] For the cluster analysis of education levels all source codes and documentations can be found in the subfolders of „\American Community Survey\(3) Cluster Analysis Education". For the cluster analysis of ages all source codes and documentation can be found in the subfolders of „\American Community Survey\(4) Cluster Analysis Age".
[192] Education levels have been defined according to their availability in the American Community Survey (see Table 17, p. 100).

Cluster analysis for age

In addition to the cluster algorithm explained above (see Part B, Sections 3.4.4.1 and 3.4.4.2), further economic considerations must be applied when age groups are clustered. First, besides mean and variance of income age must be considered as a third variable of interest. If this was not done, the algorithm would result in a clustering not sensible in economic terms. For example, the clustering of ACS wave 2000 by unweighted mean and variance of income (without considering age) to 7 clusters results in the following age clusters.

Cluster number	Ages included in that cluster
1	16-22, 92
2	23-28, 78, 82-84, 87-89, 93
3	29-35, 37
4	36, 38-41, 43, 44, 57
5	42, 45-56, 58-61
6	62-68
7	69-77, 79-81, 85, 86, 94

Figure 15: Clustering of age groups where groups should only be similar concerning mean and variance of income but not age. Exemplary for ACS wave 2000; clustered over unweighted sample; separate for ACS wave 2000.
Source: Own calculations based on ACS 2000.193

In order to cluster age groups located next to each other, it is necessary to add age as third variable of interest. For reasons outlined in Part B, Section 3.4.4.1 age is (like mean and variance of income) also z-transformed.

Second, clustering over mean, variance, and age is performed separately for people in their work ages (15 to 65) and those older than full retirement age (i.e., 66 years and older=. This separation is done because pensioners show an unsystematic pattern of mean and variance of income as discussed in the descriptive analysis in Part B, Section 3.4.3. While income exhibits a downward trend after retirement, certain age groups have much higher mean incomes than older and younger pensioners (see Figure 7, p. 105 and Figure 8, p. 106). This can be due to cohort effects or the decreasing sample sizes for retired people since single outlier bias mean income. The unsystematic change in mean and variance of income can also be seen in Figure 15, p. 118. While older people tend to belong to higher age clusters (Cluster 6 and 7 in the example above) there are certain ages which break away from this pattern. For example, concerning mean and variance of income for 92 years old people seem to be more similar to very young people aged 16 to 22 than to those being in their ages (usually belonging to Cluster 7). Besides these technical arguments, it makes sense from an

[193] The corresponding file can be found under path "\American Community Survey\(4) Cluster Analysis Age\ 4_ACS_2000_example without age as variable_unweighted.xls".

economic perspective to separate people being in their work ages from those being retired if characteristic of income are of interest.

The **number of clusters** necessary to achieve a sufficiently high sample size to estimate all income parameters cannot be determined ex ante. The reason is that the sample size of a socio-economic group in a certain year and U.S. state is a function of the clusters defined. Therefore, people in their working ages (ages 15 to 65) are clustered to different numbers of clusters from three to ten clusters, while those having reached full retirement age are clustered up to two clusters.

The clustering over age groups **results in very different cluster definitions depending on the four types of clustering (see Table 18, p. 116) and the number of clusters created**. Exemplary for all results, Table 20, p. 119 shows the cluster definitions for separate and pooled clustering over all years using the weighted sample, where people in their working age are clustered to four clusters and retired people are clustered to two clusters.

Number of clusters		Separate clustering in each calendar year										Pooled clusters over all calendar years
		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	
Working ages	1	15-23	15-24	15-24	15-24	15-24	15-22	15-24	15-24	15-24	15-24	15-24
	2	24-32	25-32	25-36	25-34	25-31	23-43	25-32	25-30	25-36	25-34	25-34
	3	33-45	33-44	37-45	35-58	32-44	35-46	33-46	31-45	37-44	35-46	35-45
	4	46-65	45-65	46-65	59-65	45-65	47-65	47-65	46-65	45-65	47-65	46-65
Retiree	1	66-74	66-73	66-75	66-72	66-74	66-79	66-79	66-74	66-81	66-78	66-74
	2	75+	74+	76+	73+	75+	80+	80+	75+	82+	79+	74+

Table 20: Results of the cluster analysis over age exemplary for weighted clusters separate each year and pooled cluster definitions over all years.
Source: Own calculations. All results can be found in „\American Community Survey\(4) Cluster Analysis Age\weighted\Ergebnisübersicht_Weighted Clusteranalyse_Age_2000-2009.xlsx".

Table 20, p. 119 shows that the results of the separate clustering differ between certain years. While the differences do not seem to be too large at first sight (e.g., Cluster 1 for people in their working ages being defined to include people being 15-23 years old versus people being 15-24 years old), they result in considerably different estimations of income parameters when sample sizes in each cluster go down. However, relatively small sample sizes appear quite often since people are not only grouped into age clusters but also into their socio-economic groups. Recall that socio-economic groups are defined by gender, age, and education and are further differentiated by U.S. state and moving year.

Because income estimation results might be sensitive to the type of clustering used, all four types will be applied in my analysis to be able to run robustness check later on. The four types of clustering have been performed for different predetermined numbers of clusters, i.e., education has been clustered to four and five clusters, people in their work ages to three to ten clusters. How many clusters are needed to reach the required sample size is decided in the next section.

3.4.4.4 Deciding about final numbers of clusters and sample sizes

After the different education and age clusters are defined, the next task is to decide which aggregation level (that is number of cluster) is needed to achieve a certain sample size in each socio-economic group. To model the migration decision mean and variance of income for each socio-economic group must be estimated. A reliable estimation can only be performed if a critical sample size exists. In general threshold values such as 30 are reported in the literature.

In order to find out which level of aggregation is needed to fulfill the sample size requirements, the actual sample size for all possible constellations of gender, age, and education are calculated exemplary for ACS wave 2000 (based on weighted mean and weighted variance of income, pooled cluster definitions are applied, see Table 17, p. 100). Results for ACS wave 2000 are illustrated and discussed in this section because it is the wave with the smallest sample size, thus also being the most critical one concerning sample size requirements. The pooled cluster definition based on weighted income measures in Table 17, p. 100 is arbitrarily chosen.

The lowest aggregation level tested here is the combination of five education clusters, six age clusters for people aged 15 to 65, and two age clusters for pensioners – separated by gender. Lower levels of aggregation on age and education are not tested because a look at the cluster definitions already reveals that the sample size requirements will not be met. Table 21, p. 121 gives an overview of the resulting socio-economic groups. Those socio-economic groups that have at least a sample size of 30 in all U.S. states in 2000 are labeled “ok”.

Age Cluster		Education Cluster		Gender	
				Male	Female
1	15-24 years old	1	Less than high school	ok	ok
		2	High school graduate	ok	
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
2	25-30 years old	1	Less than high school		
		2	High school graduate		
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
3	31-34 years old	1	Less than high school		
		2	High school graduate		
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
4	35-45 years old	1	Less than high school		
		2	High school graduate	ok	ok
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
5	46-61 years old	1	Less than high school		
		2	High school graduate	ok	ok
		3	Associate's degree		ok
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
6	62-65 years old	1	Less than high school		
		2	High school graduate		
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
7	66-74 years old	1	Less than high school		
		2	High school graduate		
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		
8	75 or older	1	Less than high school		
		2	High school graduate		
		3	Associate's degree		
		4	Bachelor’s degree		
		5	Master/Professional or higher degree		

Table 21: Overview of socio-economic groups that meet the sample size requirements of at least 30 in all 51 U.S. states (labeled “ok”) exemplary for ACS wave 2000.
Source: Own calculations based on ACS wave 2000.[194]
Empty cells have a sample size smaller than 30. Socio-economic group definitions resulting from the lowest aggregation level (8 age, 5 education and 2 gender categories). Cluster definitions resulting from weighted sample; pooled cluster definitions applied.

Table 21, p. 121 includes 80 socio-economic groups (resulting from five education cluster, eight age cluster and two gender categories) for which the income parameters must be estimated in each of

[194] The Excel file with the corresponding raw data and this table can be found under path „\American Community Survey\(5) Final Cluster Variables for Age & Education\weighted (a,b)\(5a) Pooled, 15-65, 66+,7Age 4Edu\Sample Size for all cells ACS_2000.xlsx”. The table can be found on sheet “Table for written work 9Age5Edu”.

the 51 U.S. states in order to model the migration decision. If the sample size of a certain socio-economic group (e.g., male high school graduates being 31 to 34 years old [age cluster 3, education cluster 1]) is smaller than 30 in only a single U.S. state in 2000, the migration decision of all individuals with these socio-economic characteristics in 2000 cannot be empirically analyzed.

Obviously, the lowest aggregation level is far away from meeting the sample size requirements making it impossible to model the migration decision. In 2000 it would be only eight socio-economic categories of people that could be surveyed (labeled "ok" in Table 21, p. 121) – an unacceptable situation.

Further aggregation to only four education levels (as outlined in Table 19, p. 117) leaves the situation almost unaltered. Therefore, higher aggregation levels are tested step by step until four education levels and six age levels (four clusters for people aged 15 to 65 and two clusters for pensioners) are reached. Because this still does not solve the problem for most socio-economic groups, further aggregation is needed. First, people aged 66 and older are aggregated into one age cluster of "pensioners".

Table 22, p. 123 gives the resulting socio-economic groups where the migration decision of people belonging to a socio-economic group labeled "ok" can be modeled. All others cannot be modeled because of sample sizes being smaller than 30 in at least one U.S. state in 2000.

Age Cluster		Education Cluster		Gender	
				Male	Female
1	15-24 years old	1	Less than high school	ok	ok
		2	High school/Associate's degree	ok	ok
		3	Bachelor's degree		
		4	Master/Professional or higher degree		
2	25-34 years old	1	Less than high school		
		2	High school/Associate's degree	ok	ok
		3	Bachelor's degree		
		4	Master/Professional or higher degree		
3	35-45 years old	1	Less than high school		
		2	High school/Associate's degree	ok	ok
		3	Bachelor's degree		
		4	Master/Professional or higher degree		
4	46-65 years old	1	Less than high school		
		2	High school/Associate's degree	ok	ok
		3	Bachelor's degree		
		4	Master/Professional or higher degree		
5	66 or older "pensioners"	1	Less than high school		
		2	High school/Associate's degree		ok
		3	Bachelor's degree		
		4	Master/Professional or higher degree		

Table 22: Overview of socio-economic groups that meet the sample size requirements of at least 30 in all 51 U.S. states (labeled "ok") exemplary for ACS wave 2000. Socio-economic group definitions resulting from5 age levels, 4 education levels and 2 gender categories. Cluster definitions resulting from weighted sample; pooled cluster definitions applied. Source: Own calculations based on ACS 2000.[195]

Table 22, p. 123 shows that further aggregation is needed because the income parameters for all people having more than a high school diploma or associate's degree cannot be modeled in 2000. The situation is alike in all other ACS waves from 2000 to 2009 and all other types of clustering.[196] Although needed for sample size reasons, further aggregation using the cluster analysis along education or age does not sufficiently increase sample sizes. First, further aggregation over age does not increase sample sizes of the critical education categories (less than high school degree, Bachelor and Master/Professional) high enough. Second, further aggregation over education would even out income parameters of groups that are very different to one another (as the descriptive analysis has shown), while their distinction is economically meaningful.

3.4.5 Further aggregation by economic considerations

Deleting socio-economic groups with a sample size of less than 30 (see empty cells in Table 22, p. 123) from the analysis is no alternative since this means risk-attitudes cannot be estimated for migrants with these particular socio-economic characteristics. This would bias the results of the

---

[195] The Excel file with the corresponding raw data and this table can be found under path „\American Community Survey\(5) Final Cluster Variables for Age & Education\weighted (a,b)\(5a) Pooled, 15-65, 66+,7Age 4Edu\Sample Size for all cells ACS_2000.xlsx" on sheet "Table for written work 4Edu 5 A".

[196] See Excel Files in folder „\American Community Survey\(5) Final Cluster Variables for Age & Education" for detailed sample sizes and socio-economic group definitions.

remaining observations when the relation of risk-attitudes and socio-economic characteristics is analyzed (third research question of my study). Therefore, further aggregation by economic considerations is discussed in this section. Note, in this section again all numbers discussed result from ACS wave 2000. Arguments and problems outlined below can be applied to all other waves as well.

The first problem addressed, is the small sample size for the **youngest (under 25 years old) having a college degree** (see Table 22, p. 123). This is not surprising if it is considered that only a small fraction of all people under 25 is in the age where a bachelor's degree is at least theoretical possible. In other words, the majority of people in this age group is simply too young to have a bachelor's degree. A plausible solution is to summarize bachelor and master graduates for people under 25. This approach is also supported by the descriptive analyses where mean and variance of income of bachelor and master graduates are very close for people aged 15 to 24 (see Figure 12, p. 111). Because this aggregation still causes enormous numbers of critical cases (namely in 37 U.S. states for men and in 33 U.S. states for women)[197], it is further aggregated over gender, reducing about two thirds of these critical cells. Table 23, p. 126 gives an overview of the final aggregation and its resulting sample size in each year.

A second problem is the small sample size of men and women with education "Less than high school" at age of 25 and older. Therefore, men and women are summarized for the lowest education level for all age groups of the workforce except for those under 25. This view is also supported by Figure 9, p. 107 and Figure 10, p. 107, which shows that if any, gender can only be aggregated for the lowest education level.

Finally, the sample size for retired people is too small in almost all cases. Because I investigate economic migration, it is not pensioners but people in the workforce who are in the focus of my study. Therefore, a higher aggregation level for retired people is legitimate (in contrast to higher aggregation for people in the workforce). While men and woman are still considered separately, education levels are reduced to two, namely "education up to associate's degree" and "college degree", the latter summarizing bachelor, master and professional degrees.

[197] In the ACS sample of 2000 (cluster definitions based on weighted sample and pooled cluster defintions) the socio-economic group age 15-24, education Bachelor/Master has a sample size smaller than 30 in 37 U.S. states for men and 33 U.S. states for women. Detailed data can be found in file „\American Community Survey\(5) Final Cluster Variables for Age & Education\weighted (a,b)\(5a) Pooled, 15-65, 66+,7Age 4Edu\ Sample Size for all cells ACS_2000.xlsx" on sheet "7Age 4 Edu – Analysis", cells "BC96:BC97".

3.5 Final socio-economic groups and sample sizes[198]

To get a better overview of the final definition of socio-economic groups and the resulting sample sizes reached in each group in all years, Table 23, p. 126 gives the number of states that have a sample size smaller than 30 (shown against a darker background) based on weighted ACS samples and pooled clustering.

Table 23, p. 126 shows that there is a total of 352 cells having a sample size smaller than 30, which corresponds to only 2.3% of the 15,300 cells (30 age-education-gender combinations in U.S. 51 states over 10 years). To get an impression on how critical sample sizes are, note that if the threshold is reduced to a sample size of 20, the critical number is even reduced to 141 (0.92%), where mainly the most educated among the 25 to 34 years old people are affected by the reduction in critical cells.[199]

While the pattern of critical cells (i.e., the socio-economic groups having a critical sample size (highlighted in Table 23, p. 126)), is the same for all four types of clustering, the number of U.S. states with a sample size smaller than 30 differs a lot by type of the clustering (see Table 24, p. 127).

[198] All sample sizes and different levels of aggregation can be found in folder „\American Community Survey\(5) Final Cluster Variables for Age & Education“. The resulting definition of socio-economic groups presented in Table 23, p. 126 is applied to the individuals in all ACS sample from 2000 to 2009 in folder „\American Community Survey\(6) Socioeconomic Group Variables”.
[199] Detailed data can be found under path "\American Community Survey\(5) Final Cluster Variables for Age & Education\weighted (a,b)\(5a) Pooled, 15-65, 66+,7Age 4Edu\ Problemüberischt Fallzahl Gewichtet.xlsx“, sheet "Aggregated; Cells <20”.

			American Community Survey									
Age	**Education**	**Gender**	**´00**	**´01**	**´02**	**´03**	**´04**	**´05**	**´06**	**´07**	**´08**	**´09**
15-24 years old	**Less than high school**	**Male**	0	0	0	0	0	0	0	0	0	0
		Female	0	0	0	0	0	0	0	0	0	0
	High school/ Associate	**Male**	0	0	0	0	0	0	0	0	0	0
		Female	0	0	0	0	0	0	0	0	0	0
	Bachelor/Master and higher		22	1	1	1	0	2	1	0	0	0
25-34 years old	**Less than high school**		18	1	0	1	2	3	2	1	1	2
	High school/ Associate	**Male**	0	0	0	0	0	0	0	0	0	0
		Female	0	0	0	0	0	0	0	0	0	0
	Bachelor's degree	**Male**	7	0	0	0	0	1	0	0	0	0
		Female	3	0	0	0	0	0	0	0	0	0
	Master/Profess-ional or higher	**Male**	31	10	12	10	5	7	7	7	7	7
		Female	27	4	6	2	2	6	5	4	4	2
35-45 years old	**Less than high school**		12	0	0	0	0	1	0	1	2	1
	High school/ Associate	**Male**	0	0	0	0	0	0	0	0	0	0
		Female	0	0	0	0	0	0	0	0	0	0
	Bachelor's degree	**Male**	3	0	0	0	0	0	0	0	0	0
		Female	1	0	0	0	0	0	0	0	0	0
	Master/Profess-ional or higher	**Male**	20	0	0	0	0	1	3	3	2	3
		Female	18	1	0	0	0	1	2	2	2	0
46-65 years old	**Less than high school**		1	0	0	0	0	0	0	0	0	0
	High school/ Associate	**Male**	0	0	0	0	0	0	0	0	0	0
		Female	0	0	0	0	0	0	0	0	0	0
	Bachelor's degree	**Male**	1	0	0	0	0	0	0	0	0	0
		Female	2	0	0	0	0	0	0	0	0	0
	Master/Profess-ional or higher	**Male**	3	0	0	0	0	0	0	0	0	0
		Female	7	0	0	0	0	0	0	0	0	0
Retired 66 and older	**max. Associate**	**Male**	0	0	0	0	0	0	0	0	0	0
		Female	0	0	0	0	0	0	0	0	0	0
	Bachelor/Master and higher	**Male**	9	0	0	0	0	0	0	0	0	0
		Female	15	0	0	0	0	0	0	0	0	0

Table 23: Number of U.S. states with a sample size smaller than 30, where numbers greater than null are highlighted. Source: Own calculation based on ACS waves 2000 to 2009.[200] Pooled cluster definition based on weighted sample.

[200] The original file can be found under path „\American Community Survey\(5) Final Cluster Variables for Age & Education\weighted (a,b)\(5a) Pooled, 15-65, 66+,7Age 4Edu\ Problemüberischt Fallzahl Gewichtet.xlsx". The resulting definition of socio-economic groups is performed via SPSS-programming that can be found in folder „\American Community Survey\(6) Socioeconomic Group Variables".

	ACS sample is	
Cluster definition	**weighted**	**unweighted**
	sample size <30	
Pooled	352 (2.3%)	572 (3.7%)
Separate	642 (3.8%)	625 (4.1%)
	sample size <20	
Pooled	141 (0.9%)	314 (2.1%)
Separate	493 (2.9%)	389 (2.5%)

Table 24: Number of U.S. states summed up from 2000 to 2009 with sample size smaller than 30 by type of cluster definition (as percentage of total cells given in parentheses).
Source: Own calculation based on ACS waves 2000 to 2009.[201]

Note that the vast majority of critical cells in Table 24, p. 127 originates from ACS wave 2000 in all types of cluster definitions.

3.6 Estimating annual income parameters by neighboring waves[202]

3.6.1 Problem and solution

Problem

Even though the clustering of socio-economic groups has reduced the number of income parameters to be estimated from originally 587,520[203] to 16,320 reaching much higher sample sizes, there are still some socio-economic groups for which income parameters cannot be estimated because of low sample sizes (see Table 23, p. 126 and Table 24, p. 127). Recall, that even if only a single U.S. state of a certain socio-economic group has a sample size smaller than 30, the migration decision of all individuals with this very socio-economic characteristic cannot be empirically analyzed. For example, in Table 23, p. 126 this is the case for all men between 46 and 65 years of age(by the time of the move in 2000) with a bachelor's degree – no matter from which country they originate or migrate to.

[201] The original file can be found under path "\American Community Survey\(5) Final Cluster Variables for Age & Education\ Critical Values Weighted versus Unweighted.xlsx".
[202] All income parameters of all years are finally summarized in folder „\American Community Survey\(12) Table of all Income Measures in all years", where files separate for each wave can be found in folder „\American Community Survey\(11) Original and adjusted Income Parameters by_SocEconGr". Income parameters based only on the annual sample are can be found in folder „\American Community Survey\(7) Income Parameters added to_Individuals". Data is then compressed to one record for each socio-economic group in folder „\American Community Survey\(8) Income Parameters by_SocEconGr". Income parameters over neighboring waves of data are estimated using all individuals available in all surrounding years in folder „\American Community Survey\(9) Adjusted Income Parameters by_Individuals". Data is then compressed to one record for each socio-economic group in folder „\American Community Survey\(10) Adjusting Income Parameters by_SocEconGr".
[203] 2 gender categories, 96 age levels, 6 education levels in 51 U.S. states from 2000 to 2009 adds up to 599,040 socio-economic groups.

Sample sizes cannot be increased by further aggregation as the descriptive analysis has shown. Recall, for example, that women and men holding a master's degree or higher degree differ a lot in mean and variance of income (see Figure 9, p. 107 and Figure 10, p. 107). Therefore, other solutions must be found.

Solution

There are two competing solutions to that problem. First, migrants belonging to a socio-economic group with critical sample size are canceled out of the sample of migrants. This assures that risk-attitudes analyzed in my study are based on income parameters that are estimated based on a sample size of at least 30. At the same time applying this solution would mean that the risk-attitude of people with certain socio-economic characteristics cannot be estimated. This probably distorts the analysis of risk-attitudes by socio-economic characteristics later on.

The second solution is the merging of identical socio-economic groups of surrounding ACS waves. For example, if the sample size of the socio-economic group with men aged 25 to 34, having less than a high school diploma and living in Idaho, is too small in the ACS wave 2003, the sample size of the same socio-economic group (men aged 25 to 34, less than a high school, Idaho) from the previous year 2001 and the next year 2005 are merged with the sample of 2003. If the sample size of the first wave in 2000 is too small, it can only be merged with identical socio-economic groups in 2003; small sample sizes in 2009 can only be merged with identical socio-economic groups of 2007.

This procedure can be applied because the American Community Survey is run throughout the year which means dollar amounts do not reflect calendar year dollars but give the amount earned in the last 12 month.[204] Because of data privacy protection the exact date of the interview of each person is not given in the data. This means that income reported for a person can refer to anything from the previous calendar year (if the interview took place at the beginning of January) to the current year (if the interview took place at the end of December). As the reference periods of consecutive waves of data overlap anyway, merging the samples should not distort the results.

Usually, when dollar amounts of different years are merged, they should be made comparable by inflation adjustment. Because of the overlapping reference periods the editor of the American

[204] See Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable „INCTOT", paragraph on "Description".

Community Survey data set, the Minnesota Population Center at the University of Minnesota, itself does not recommend to account for inflation.[205]

Since it is not clear whether an aggregation with or without inflation adjustment is the right approach, I decide to apply both competing solutions and run robustness checks later on. The inflation adjustment is performed using the percentage change of the annual average of the Historical Consumer Price Index for All Urban Consumers in U.S.A.[206]

3.6.2 SPSS-implementation: estimation over neighboring waves

Problem: Personal weights of different waves cannot be compared[207]

Aggregating samples of identical socio-economic groups over neighboring waves is without any problem when samples are unweighted. But once weights are applied, this can result in considerable distortions. The reason is that weights of a specific year are calculated to result in representative samples for this specific year. This means different weights can be applied to one and the same person in different years depending on the structure of the U.S. population and sampling characteristics in that specific year.[208]

Applying original weights without adjustment to the aggregated sample would distort estimations in an unpredictable way. A stylized example may clarify the problem. Table 25, p. 130 gives an example of 6 individuals aggregated over 3 waves of data with 2 individuals originating from each year. Note that within each year the first individual has exactly half the weight of the second individual. To produce estimates like mean income within a certain year, each individual's income must be weighted.[209] In this case it is the relation of weights between the first and the second individual in each year that counts rather than the absolute amount.

---

[205] See Minnesota Population Center, University of Minnesota [No date b].
[206] See U.S. Bureau of Labour Statistics (2001), Table 24, p. 73.
[207] Weights are adjusted separate for each year in folder „\American Community Survey\(9) Adjusted Income Parameters by_Individuals\(9b) income raw, inflation adjusted and weight adjusted".
[208] „Weights are used to bring the characteristics of the sample more into agreement with those of the full population by compensating for differences in sampling rates across areas, differences between the full sample and the interviewed sample, and differences between the sample and independent estimates of basic demographic characteristics" (U.S. Census Bureau (2009c), pp. 11-1). For more details about how weights are applied in the American Community Survey see U.S. Census Bureau (2009c), pp. 11-1 to 11-20.
[209] See Minnesota Population Center, University of Minnesota [No date d].

Year person was surveyed	Original personal weights	Adjusted personal weights
2001	100	1/3
2001	200	2/3
2002	50	5/15
2002	100	10/15
2003	200	2/6
2003	400	4/6

Table 25: Example of original and adjusted personal weights when data of neighboring waves is aggregated. Source: Own tabulation.

If the original weights are applied to the aggregated sample, this would mean the first person (first row in Table 21, p. 119) is considered to have double the weight of the third person (third row in Table 21, p. 119). This would be a misinterpretation as weights reported in the American Community Survey do not contain any information about the relation of records of one wave to any other wave. Weights only give the relational weights of records within the same year.

Solution: Weights adjustment

In order to make records of different waves comparable but still keeping their rank within each year, weights of each year are adjusted to 1. This approach ensures that records of each year are weighted equally. Adjusted weights for the example are also given in Table 25, p. 130.

4 Creating the final data set

The last step of data cleaning is to merge all data sets to create the final data set from which risk-attitudes are finally estimated. This section gives an overview of the cleaned data sets that are later on merged, explains how they are merged, and finally presents the structure of the final data sets.

4.1 Stylized structure of cleaned PSID and ACS data

The data cleaning of the Panel Study of Income Dynamics results in a cleaned data set for each migrant wave which includes all migrants of that particular wave with their socio-economic characteristics gender, age, and education at the time of the move. Due to the difficulties in defining education at the time of the move (see Part B, Section 2.3.3.6) each migrant has six different education definitions (see Table 14, p. 82). Additionally, further variables are included in the data set, i.e., state of origin and destination, year of the move, the Move Context Variable (Part B, Section 2.4), all Family Interview Numbers and Sequence Numbers from 1968 to 2009 and all other variables included in the Cross-Year Individual File from 1968 to 2009. Figure 16, p. 132 gives the stylized structure of the cleaned data set of the Panel Study of Income Dynamics.

The cleaning of the American Community Survey results in a single data set that includes all annual income parameters for all socio-economic groups and the sample sizes from which the income parameters were estimated. Figure 17, p. 132 gives a stylized overview of the respective data set.

						6 education definitions							All Interview and Sequence Numbers 1968 -2009					All variables of the Cross-Year Individual File								
	State of origin	Destination state	Year moved	Age moved	Gender	…	…	…	…	…	…	Move Context	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1																										
2																										
…																										

Figure 16: Stylized structure of cleaned data of the Panel Study of Income Dynamics.
Source: Own illustration based on own calculations. Such a file exists separate for each migrant wave.

	36 annual income parameters						8 sample sizes N on specifications						8 flag variables indicating sample size					
Socio-economic group [AEG/SS/YY]*	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1																		
…																		
15,300																		

Figure 17: Stylized structure of cleaned ACS data.
*** Socio-economic groups are coded [AEG/SS/YY] which means: one digit for the age cluster [A], one digit for the education level [E], one digit for gender [G], two digits for the U.S. state and two digits for the year of the move [YY].**
Source: Own illustration based on own calculations.[210]

[210] Final file resulting from ACS data can be found in folder „\American Community Survey\(12) Table of all Income Measures in all years\12_All Income Measures in all years.sav".

The variable on socio-economic groups in the American Community Survey is coded as follows: The first digit indicates the age cluster [A], the second digit indicates the education level [E], and the third digit indicates gender [G]. The U.S. state is indicated as [SS] by the fourth and fifth digit in the socio-economic variable. The year of the move [YY] is indicated by the last two digits of the variable.

In each year and for each of the U.S. 51 states there are altogether 30 combinations of gender, age, and education (see column on Gender [G] in Table 26, p. 133) The resulting 1,530 combinations (51 U.S. states multiplied by 30 gender-age-education combinations) must be estimated for each of the 10 years surveyed, which sums up to all together 15,300 socio-economic groups in my analysis. They result from the clustering performed in Part B, Section 3.4. Table 26, p. 133 repeats all combination of gender, age, and education (see also Table 23, p. 126).

Age [A]	Education [E]	Gender [G]
15-24 years old	Less than high school	Male
		Female
	High school/ Associate	Male
		Female
	Bachelor/Master and higher	
25-34 years old	Less than high school	
	High school/ Associate	Male
		Female
	Bachelor's degree	Male
		Female
	Master/Professional or higher	Male
		Female
35-45 years old	Less than high school	
	High school/ Associate	Male
		Female
	Bachelor's degree	Male
		Female
	Master/ Professional or higher	Male
		Female
46-65 years old	Less than high school	
	High school/ Associate	Male
		Female
	Bachelor's degree	Male
		Female
	Master/Professional or higher	Male
		Female
Retired 66 and older	max. Associate	Male
		Female
	Bachelor/Master and higher	Male
		Female

Table 26: 30 final socio-economic groups clustered for gender, age, and education which exist in each of the 51 U.S. states over 10 years.
Source: Own illustration based on own calculations. Same definitions of socio-economic groups like in Table 23, p. 126. Socio-economic groups where gender or education has been aggregated are shown against a darker background.

The 36 annual income parameters in Figure 17, p. 132 result from originally three parameters on income (namely mean, variance, and semi-variance) which are estimated for four different cluster definitions (namely weighted/unweighted and pooled/separate, see Table 18, p. 116) and by three different samples (namely the annual sample, the a aggregated sample of surrounding waves adjusted only by weights and adjusted by weights and inflation). Three parameters multiplied by four cluster definitions multiplied by three samples results in all together 36 income parameters.

Sample sizes only vary for the four cluster definitions and between two sample compositions resulting from annual data only and from data aggregated by surrounding waves. The sample size is the same no matter which of the three parameters is estimated or whether data of surrounding waves is adjusted only by weights or also by inflation. Therefore, only eight variables (four cluster definition multiplied by two sample compositions) on sample sizes are needed.

4.2 Adding socio-economic group variables to cleaned PSID data[211]

<u>Problem</u>

For the empirical analysis migrants (identified in the Panel Study of Income Dynamics) must be related to their specific income parameters (estimated from the American Community Survey) depending on their socio-economic characteristics. Technically speaking, a single variable on socio-economic characteristics is needed which relates migrants to income parameters. So far such a variable only exists in the American Community Survey, while the Panel Study of Income Dynamics includes variables separate for gender, age, education, and year of the move (see Figure 16, p. 132 and Figure 17, p. 132).

<u>Solution</u>

A variable is added to the Panel Study of Income Dynamics which summarizes information about gender, age, education, and the year of the move in a single variable. In contrast to the socio-economic group variable in the American Community Survey which is coded [AEG/SS/YY], the variable added to the Panel Study of Income Dynamics is coded [AEG/YY]. It does not include two digits on the U.S. state [SS] because the estimation of risk-attitudes is based on the comparison of income possibilities a migrant faces in all 51 U.S. states. Once migrants are defined into socio-economic groups by gender, age, and education (see Table 26, p. 133), and information about the

[211] The resulting SPSS data file exemplary for those moving in 2000 is „\Mover\(3) Add SocEcon Group\3_SocEcon_No State_2000.sav". It is created by first adding cluster variables on age and education that are the same in all years. This is done in the SPSS-Syntax file „\Mover\(1) Mover between Waves\Collect all mover, add EduCl and AgeCl Pool.sps". Then age cluster variables which are different year by year are added to each year separately in the SPSS Syntax file „\Mover\(2) Add Age Cluster for Separate\(2) Add Age Cluster for Separate_2000.sps". Finally, corresponding Socio-Economic Group variables are added based on the cluster variables on education and age. This is done in the SPSS Syntax file „\Mover\(3) Add SocEcon Group\(3) Add SocEcon_No State_2000.sps".

year of the move is added, income parameters of all U.S. states can be compared for all people belonging to the same combination of gender, age, education, and year.

Because of the six different definitions of education in the Panel Study of Income Dynamics (see Table 14, p. 82) and the four different cluster definitions in the American Community Survey (namely weighted/unweighted and pooled/separate, see Table 18, p. 116) it is not enough to add a single variable on socio-economic characteristics to the Panel Study of Income Dynamics. Instead, 6 education definitions multiplied by four types of clustering results in 24 different definitions of the socio-economic group. Thus, 24 variables on socio-economic characteristics must be added for each migrant in the Panel Study of Income Dynamics. Figure 18, p. 135 illustrates all 24 combinations and gives the names of the variables in the resulting data set.

4 types of clustering in the ACS (see Table 18**, p. 116)**	**6 education definition in the PSID (see** Table 14**, p. 82)**		**24 socio-economic group variables for each migrants must be added**
Pooled clustering over all years by weighted income parameters	Education 1 … Education 6		STD_SocEcon_WPool_NoSt1 … STD_SocEcon_WPool_NoSt6
Pooled clustering over all years by unweighted income parameters	Education 1 … Education 6	⇒	STD_SocEcon_UNwPool_NoSt1 … STD_SocEcon_UNwPool_NoSt6
Separate clustering for each year by weighted income parameters	Education 1 … Education 6		STD_SocEcon_WSep_NoSt1 … STD_SocEcon_WSep_NoSt6
Separate clustering for each year by unweighted income parameters	Education 1 … Education 6		STD_SocEcon_UNwSep_NoSt1 … STD_SocEcon_UNwSep_NoSt6

Figure 18: 24 socio-economic group variables for each migrant in the Panel Study of Income Dynamics.
Source: Own illustration based on own definitions.

Each of the 24 socio-economic group variables defines migrants into socio-economic groups by gender, age, education and year [AEG/YY]. Based on this definition, income parameters for all 51 U.S. states can be taken from annual income estimations in the American Community Survey.

4.3 Deriving risk-free discount factors

In order to calculate the present value of income over the remaining lifetime and working time (as derived in Part A, Section 2.4.2), risk-free discount factors are needed. Depending on (i) the time of the move and (ii) the time period over which income should be discounted, risk-free discount factors vary for each migrant. Hence, they are migrant-specific and are therefore added to the Panel Study of Income Dynamics.

Risk-free discount factors are calculated in three steps: First, the planning period must be determined by means of adding a variable on the remaining time until reaching full retirement age and another variable on the remaining life expectancy to the Panel Study of Income Dynamics (see Part B, Section 4.3.1). Second, a table of risk-free discount factors for all years surveyed and over all possible time periods is estimated (see Part B, Section 4.3.2). Third, two risk-free discount factors are added to each migrant in the Panel Study of Income Dynamics specific to the individual time periods (remaining time until reaching full retirement age and life expectancy) and the year the move took place (see Part B, Section 4.3.3).

4.3.1 Adding full retirement age[212] and life expectancy[213]

4.3.1.1 General approach

While life expectancy is taken from the United States National Vital Statistics Reports issued by the National Center for Health Statistics[214], the remaining years until reaching full retirement age are taken from the U.S. Social Security Administration[215]. The latter can be implemented straight forward as the full retirement age depends on the year a person is born. It start with those born in 1937 or earlier which reach full retirement age with 65 years and ends with those being born in 1960 or later which reach full retirement age with 67 years.[216]

---

[212] For legal retirement ages depending on the year born see U.S. Social Security Administration (2011). The corresponding SPSS programming can be found under „\Mover\(5) Personal and Family Data Sets\ Add variables full retirement age, life expectancy and disount rates for retirement and life.sps".

[213] Data on life expectancy is collected in the excel file "\Data Supplements\Average remaining lifetime in the United States 2000-2009_Detailed.xlsx" in table „Summary".

[214] For life expectancy in 2000 see Arias, Elizabeth (2002), pp. 9-12. For life expectancy in 2001 see Arias, Elizabeth (2004a), pp. 9-12. For life expectancy in 2002 see Arias, Elizabeth (2004a), pp. 9-12. For life expectancy in 2003 see Arias, Elizabeth (2006), pp. 10-13. For life expectancy in 2004 see Arias, Elizabeth (2007), pp. 10-13. For life expectancy in 2005 see Arias, Elizabeth, Brian L. Rostron, and Betzaida Tejada-Vera (2010), pp. 10-13. For life expectancy in 2006 see Arias, Elizabeth (2010), pp. 9-12. For life expectancy in 2007 see Arias, Elizabeth (2011), pp. 10-13. For life expectancy in 2008 see Miniño, Arialdi M., Sherry L. Murphy, Jiaquan Xu, and Kenneth D. Kochanek (2011), p. 74, Table 6.

[215] For legal retirement ages depending on the year born see U.S. Social Security Administration (2011).

[216] See U.S. Social Security Administration (2011).

4.3.1.2 Data cleaning of life expectancy

Problem: Incomplete data on life expectancy for years 2008 and 2009

While the United States National Vital Statistics Reports includes data on life expectancy in detail for all ages from 0 to 100 and genders in the years 2000 to 2007, data for 2008 and 2009 is only available for selected ages separate for male and female citizens. Table 27, p. 137 illustrates the data structure exemplary for males aged 15 to 30.

	Raw data: Different data structure in 2008 and 2009										Solution	
Age	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2008**	**2009**
15	59.9	60.2	60.3	60.5	61.0	60.6	60.9	61.1	61.3	61.5	61.3	61.5
16	59.0	59.3	59.4	59.6	60.0	59.7	59.9	60.2			60.4	60.5
17	58.0	58.3	58.4	58.6	59.0	58.7	59.0	59.2			59.4	59.6
18	57.1	57.4	57.5	57.7	58.1	57.8	58.0	58.3			58.5	58.6
19	56.2	56.4	56.5	56.7	57.2	56.8	57.1	57.3			57.5	57.7
20	55.2	55.5	55.6	55.8	56.2	55.9	56.1	56.4	56.6	56.7	56.6	56.7
21	54.3	54.6	54.7	54.9	55.3	55.0	55.2	55.5			55.7	55.8
22	53.4	53.7	53.8	54.0	54.4	54.1	54.3	54.5			54.8	54.8
23	52.4	52.7	52.8	53.0	53.4	53.1	53.4	53.6			53.8	53.9
24	51.5	51.8	51.9	52.1	52.5	52.2	52.5	52.7			52.9	52.9
25	50.6	50.9	51.0	51.2	51.6	51.3	51.5	51.8	52.0	52.0	52.0	52.0
26	49.7	49.9	50.1	50.3	50.7	50.4	50.6	50.9			51.1	51.1
27	48.7	49.0	49.1	49.3	49.7	49.4	49.7	49.9			50.1	50.1
28	47.8	48.1	48.2	48.4	48.8	48.5	48.8	49.0			49.2	49.2
29	46.9	47.1	47.3	47.4	47.9	47.6	47.8	48.1			48.2	48.2
30	45.9	46.2	46.3	46.5	46.9	46.6	46.9	47.1	47.3	47.3	47.3	47.3

Table 27: Life expectancy for males ages 15 to 30 in the U.S.A.
Only the first decimal of life expectancy is given here, while the correct number with all decimals is used for the calculations. Ages with missing data in 2008 and 2009 are reported against a darker background.
Source: Own illustration based on Arias (2002, 2004a, 2004b, 2006, 2007, 2010), Arias, Rostron, and Tejada-Vera (2010), Miniño, Murphy, Xu, and Kochanek (2011), Kochanek, Xu, Murphy, Miniño, and Kung (2011).

For years 2008 and 2009 life expectancy of males aged 15 to 30 is only given for males being exactly 15, 20, 25, and 30 years old. Because detailed data on life expectancy for all age groups is missing for 2008 and 2009, it must be decided whether data for years 2000 - 2007 is reduced to the selected age groups given in 2008/09 or missing data of 2008/09 is approximated.

Solution

Missing data for 2008/09 is approximated by linear interpolation- This means, when age rises by one year, life expectancy is falling by one fifth of the difference between two given life expectancies since there are 5 years between ages for which data is given. For example, the difference of life expectancy for males aged 15 and 20 in 2008 is 4.7. Approximated life expectancy for those aged 16 in 2008 thus is 61.3 minus one fifth of 4.7 equals 60.4. Note that Table 27, p. 137 reports only the first decimal for years of life expectancy, while the exact number of years is taken for further calculations.

Linear interpolation is chosen for two reasons: First, detailed data for years 2000 to 2007 is not lost. Second, life expectancy estimated for 2008/09 is consistent with data of years 2000 to 2007: in each year of data life expectancy is monotonically decreasing by age; with exception of the year 2005 and some singular cases, each age level has a trend of slightly rising life expectancy over the years 2000 to 2009.

4.3.2 Table of risk-free discount factors for all years and time periods[217]

Problem 1: Risk-free term-structure

The theoretical migration decision model developed in Part A needs risk-free discount factors. Theoretically, they can easily be determined from the term structure of interest rates. Unfortunately, neither the U.S. Department of the Treasury nor the U.S. Federal Reserve or the Central Reserve Bank of any other U.S. state publishes the term structure of interest rates for the U.S. market.[218]

Solution to Problem 1

Similar to the approach taken by the Deutsche Bundesbank[219] and many other central banks[220] the U.S. Federal Reserve Board estimates the treasury yield curves[221] with the extension of the Nelson-Siegel (1987) approach proposed by Svensson (1994).[222] These estimations are available as continuously updated data attachments to the paper of Gurkaynak, Sack, and Wright (2006) continuously compounded for maturities from 1 to 30 years on a daily basis from 1961 to

[217] Risk-free discount factors for all years and maturities including raw data can be found in the excel file "\Data Supplements\Gurkaynak et al. (2006) - The US Treasuy Yield Curve, 1961 to present –DATA.xlsx".
[218] See for an announcement concerning the U.S. Department of the Treasury U.S. Department of the Treasury (2011b), p. 1. Concerning the Federal Reserve Bank and all other Central Reserve Bank of U.S. states the statement is due to author's own research.
[219] See Schich (1997).
[220] The term structure is also estimated with the extension of the Nelson-Siegel (1987) approach proposed by Svensson (1994) by the National Bank of Belgium, the Bank of France, the Bank of Italy, the Central Bank of Norway, the Bank of Spain and Swiss National Bank (see Bank for International Settlement (2005), pp. 1, 8, 12, 20, 24, 28).
[221] In the context of the U.S. market the term structure is usually called yield curve (see U.S. Department of the Treasury (2011a)).
[222] For the procedure to estimate the term structure by the Deutsche Bundesbank see Deutsche Bundesbank (1997), p. 64. For the procedure to estimate the yield curve by the U.S. Federal Reserve Board see Gurkaynak, Sack, and Wright (2006), p. 13.

February 2012.[223] They also publish the parameter estimates from which discount factors over periods longer than 30 years can easily be derived.[224]

Problem 2: Exact moving date is unknown

The term structure can be determined on a daily basis and hence might change from day to day. In contrast, the exact date of the move is unknown. A variable on month of the move exists, but is often not specific enough: Besides the exact month, respondents can only give the season the move took place (winter, spring, summer, fall) or they can even state they do not know it anymore.[225] Because the month of the move is unknown for many migrants, the only thing being available for all migrants is the year of the move. Since the move could have been taking place during all days of the year, it is not clear which day's term structure is relevant for the move. This means, the frequency of migration data (annual) and data on term structure (daily) is not congruent.

Solution to Problem 3

The annual term structure is approximated by simply taking the average of daily term structures of the respective year. Note that this approach is in line with the way income is reported in the American Community Survey where annual income includes income of the last 12 month.[226] Because interviews are run throughout the year, annual income usually does not reflect calendar year dollar amounts depending on the time of the interview.[227]

[223] See data set attached to Gurkaynak, Sack, and Wright (2006).
[224] See Gurkaynak, Sack, and Wright (2006), p. 15.
[225] See Institute for Social Research, Survey Research Center, University of Michigan (2008f), p. 23.
[226] See Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable „INCTOT", paragraph on "Description".
[227] See Minnesota Population Center, University of Minnesota [No date c], online dictionary on variable „INCTOT", paragraph on "Description".

4.3.3 Adding migrant-specific risk-free discount factors[228]

Depending on the time of the move, the remaining years until reaching full retirement age, and the remaining life expectancy, risk-free discount factors to calculate the present value of working life income and life income are added to each migrant on the Panel Study of Income Dynamics.

4.4 Stylized structure of merged data from which risk-attitudes are esimted[229]

So far, only isolated data sets have been discussed. Their information finally has to be combined in order to be able to estimate-risk-attitudes. This section gives an overview of the structure of the final data set.

Recall that for each migrant income parameters are estimated in 84 different ways. Table 28, p. 141 gives an overview. The 84 different income parameters result from different types of clustering (Column (1), Table 28, p. 141), different sample compositions (Column (2), Table 28, p. 141), and different education definitions (Column (3) Table 28, p. 141). In Table 28, p. 141 letters given in brackets and printed in italics constitute the file name of the respective constellations, e.g., WPoo-Edu1_Adj.

[228] All data is merged in the Excel file „\Mover\(5) Personal and Family Data Sets\ 5_Individual and Family Mover.xlsx" on sheets "5_Family Data Set" for all moving family members and on sheet "5_Personal Data Set" for moving Heads. The results are copied to the SPSS files in folder „\Mover". Input data in the Excel file is copied from other sources as follows: Full retirement age is copied from the SPSS data file in folder „\Mover\(5) Personal and Family Data Sets". Life expectancy is copied from the Excel file "\Data Supplements\Average remaining lifetime in the United States 2000-2009_Detailed.xlsx". The table of risk-free discount factors is copied from "\Data Supplements\Gurkaynak et al. (2006) - The US Treasuy Yield Curve, 1961 to present – DATA.xlsx".

[229] Data sets for the personal migration decision including personal income parameters can be found in folder "\Alpha\(1) Personal Alpha aus Maple\(1) Datenmasken erstellen_xlsx". Data sets for family migration decisions including family income can be found in folder „\Alpha\(2) Family Alpha aus Maple\(3) Datenmasken relevante Familieneinkommmen Head, Wife".

Types of clustering[230]	**Composition of the total sample**[231]	**Education Definition**[232]
• Pooled clustering over all years by weighted income parameters *[WPool]* • Pooled clustering over all years by unweighted income parameters *[UNwPool]* • Separate clustering for each year by weighted income parameters *[WSep]* • Separate clustering for each year by unweighted income parameters *[UNwSep]*	• Annual sample *[_]* • Neighboring waves (adjusted by income and inflation) *[_Adj]* • Neighboring waves (adjusted only by weights) *[Adj2]*	• Education 1 *[Edu1]* • Education 2 *[Edu2]* • Education 3 *[Edu3]* • Education 4 *[Edu4]* • Education 5 *[Edu5]* • Education 6 *[Edu6]*
Family migration decision: 12 files with [Edu1]		
Personal migration decision: 72 files		

Table 28: 84 different data files to finally estimate the risk-attitude for personal and family migration decisions. Source: Own illustration based on own definitions.

Since different income parameters result in different estimations of risk-attitude, I create 84 different data sets. Note that all data files include the same 321 decision-makers[233] while only the respective income parameters may differ from file to file. Figure 18, p. 135 gives the stylized structure of such a data set exemplary for migration decisions based on Head's individual income. The only difference to data sets relating to migration decision based on family income is that income parameters relate to family income rather than Head's personal income.

[230] Reasons of the four types of clustering are discussed in detail in Part B, Section 3.4.4.2.
[231] Reasons and algorithms of the three ways to create the total sample from which income parameters are estimated are discussed in detail in Part B, Section 3.6.
[232] Reasons and algorithms of the six education definitions are discussed in detail in Part B, Section 2.3.3.6.
[233] Not that although each data file initially includes 321 migrants the number may be reduced later on. Reasons for the reduction are, for example, migration decisions where it is not clear who is the second main earner (usually wife) in the family. In this case the migration decision based on family income cannot be estimated. Other reasons are discussed in Part C.

Characteristics of the decision-maker (Head)

For each planning period all income parameters are given (mean, variance etc.)

		[AEGYY]			**[AEG/State/YY]**	**Annual income**				**Present value of working life income**			
Personal MoverID	**Family MoverID**	**Socio-economic**	**Possible destinations**	**State**	**Socio-economic**	**Mean**	**…**	**…**	**…**	**Mean**	**…**	**…**	**…**
1	1	44100	12	Illinois	441**12**00								
			1	Alabama	441**01**00								
Origin	**Destination**		2	Arizona	441**02**00								
14	12		3	Arkansas	441**03**00								
			4	California	441**04**00								
Discount Factor Retire	**Discount Factor Life**		5	Colorado	441**05**00								
12.191120399	14.589363439		6	Connecticut	441**06**00								
			…	…	441**00								
			…	…	…								
			51	Hawaii	441**51**00								
…	…	…	…	…	…	…	…	…	…	…	…	…	…
		[AEGYY]			**[AEG/State/YY]**	**Annual income**				**Present value of working life income**			
Personal MoverID	**Family MoverID**	**Socio-economic**	**Possible destinations**	**State**	**Socio-economic**	**Mean**	**…**	**…**	**…**	**Mean**	**…**	**…**	**…**
3	2	32100	10	Georgia	321**10**00								
…	…	…	…	…	…	…	…	…	…	…	…	…	…

Figure 19: Stylized structure of merged data set.

Source: Own illustration based on own calculations. Data sets for the personal migration decision including personal income parameters can be found in folder "\Alpha\(1) Personal Alpha aus Maple\(1) Datenmasken erstellen_xlsx". Data sets for family migration decisions including family income can be found in folder „\Alpha\(2) Family Alpha aus Maple\(3) Datenmasken relevante Familieneinkommmen Head, Wife".

As indicated in Figure 19, p. 142, each data set includes all decision-makers listed one after the next. The first three columns of Figure 19, p. 142 always include characteristics of the decision-maker as follows:

- Unique identification of Head (PersonalMoverID) and the family he is associated with (FamilyMoverID)
- The PSID code[234] of the state of origin and destination
- Personal risk-free discount factors for working life income and lifetime income
- Socio-economic group [AEGYY] indicating the age cluster [A], education level [E], gender [G] and year of the move [YY].

In Figure 19, p. 142 the first Head has a personal and family mover ID of 1. The socio-economic group code reveals that Head belongs to the fourth age cluster (first digit), the fourth education level (second digit), is male (third digit) and has moved in 2000 (last two digits) from Iowa to Illinois (PSID-codes 14 and 12, respectively). Recall, that the fourth age cluster this Head belongs to relates to different age groups and education level depending on the 84 ways income parameters are estimated.

The fourth and fifth columns of Figure 19, p. 142 give all possible destination states starting with the actual destination state (Illinois). Note that there are 51 possible destination states since the Panel Study of Income Dynamics from which migrants are identified differentiates between 50 U.S. states and Washington D.C. In the sixth column of Figure 19, p. 142 full socio-economic group codes are generated for all 51 possible destination states. Digits four and five in the full socio-economic group code give the PSID state code in two digits (bold type in Figure 19, p. 142).

The final data set from which risk-attitudes are estimated results from combining (i) cleaned data of the Panel Study of Income Dynamics (see Figure 16, p. 132), (ii) cleaned data of the American Community Survey (see Figure 17, p. 132), (iii) variables on socio-economic groups relating both data sets (see Part B, Section 4.2), and (iv) migrant-specific risk-free discount factors (see Part B, Section 4.3) from which income parameters for longer planning periods are derived.

Given the full socio-economic group code [AEGSSYY] in column six of Figure 19, p. 142 all income parameters on all planning periods (annual income, present value of working life income, and lifetime income) can be merged.

[234] The PSID state codes can be found in Table 15, p. 99.

Part C – Empirical Analysis

Part C combines the theoretical model of Part A with the data set of Part B. To be more precise, Part C, Section 1 empirically estimates individual migrants' risk-attitudes in the context of migration (first research question of my dissertation). Part C, Section 3 answers the question on whether risk actually plays a significant role in the migration decision (second research question of my dissertation). Part C, Section 4 investigates the relation of socio-economic characteristics and risk-attitudes in the context of economic migration (third research question of my dissertation). Finally, Part C, Section 5 gives a conclusion.

1 Estimating individuals' risk-attitudes

1.1 Problem and solution in estimating risk-attitudes

Based on the migration decision model of Part A, an individual's risk-attitudes can be derived from a replication of the migration decision observable as follows: Since the migrant chooses the destination state that maximizes his preference value, the preference value of all other states must be smaller or equal to the preference value of the actual destination state. This can be formalized in a system of 50 linear inequalities as follows[235]

$$\begin{gathered}\left[E_t\left(inc_{i,d,t,T}\right)-\alpha_{i,t,T}var_t\left(inc_{i,d,t,T}\right)\right]-\left[E_t\left(inc_{i,j=1,t,T}\right)-\alpha_{i,t,T}var_t\left(inc_{i,j=1,t,T}\right)\right]\geq 0\\ \vdots\\ \underbrace{\left[E_t\left(inc_{i,d,t,T}\right)-\alpha_{i,t,T}var_t\left(inc_{i,d,t,T}\right)\right]}_{\psi_{i,d,t,T}\left(\alpha_{i,t,T}\right)}-\underbrace{\left[E_t\left(inc_{i,j=50,t,T}\right)-\alpha_{i,t,T}var_t\left(inc_{i,j=50,t,T}\right)\right]}_{\psi_{i,d,t,T}\left(\alpha_{i,t,T}\right)}\geq 0\end{gathered}\tag{12}$$

where d denotes the destination state actually chosen, $\psi_{i,d,t,T}\left(\alpha_{i,t,T}\right)$ the preference value of individual i in destination state d concerning the migration decision taken at time t with a planning period of $(T-t)$ as function of the respective individual's risk-attitude $\alpha_{i,t,T}$.

Individual's risk-attitude can then be derived by solving the system of linear inequalities for $\alpha_{i,t,T}$. In other words, the question of the theoretical model of Part A is reversed in the sense that the variable of interest is not the destination state anymore but individual's risk-attitude $\alpha_{i,t,T}$.

[235] Recall that there are 51 potential destination states (50 U.S. states and Washington D.C.) of which one state is actually chosen as destination state.

Problem: Migration decision cannot be replicated 100%

Unfortunately, the system of linear inequalities can only be solved for a small number of migration decisions in my data set. More precisely, depending on the way risk-attitude is estimated only 3% to 7% of all migration decisions can perfectly be reproduced this way.

The problem that the true migration decision cannot be replicated for most migrants is not surprising for two reasons: First, from a technical point of view a system of linear inequalities will have no solution if it contains mutually exclusive inequalities. Since the systems of linear inequalities in my study consist of 50 inequalities, it is not surprising that many of them have no solution.

Second, from an economic point of view it is also not surprising that many migration decisions cannot be replicated 100% because the migration decision model relies solely on monetary determinants. Monetary determinants are probably the main drivers of economic migration but might not be the only determinants at the same time. Recall, for example, that psychic migrations costs are neglected in my analysis.

This leaves the problem of how to estimate individual's risk-attitude from the migration decision observable.

Solution: Minimizing predictive errors

Since the migration decision cannot be perfectly reproduced in a system of 50 inequalities, the parameter for risk-attitude is calibrated to the true migration decision by minimizing predictive errors.

1.2 Methodology: non-linear minimization of predictive errors

1.2.1 When does a predictive error occur?

For any given value of risk-attitude $\alpha_{i,t,T}$, a predictive error occurs when the preference value of any non-chosen destination state is higher than the preference value of the actually chosen destination state. Vice versa, there will be no predictive error if the destination state actually chosen by the migrant is the one with the highest or equal to the highest preference value among all potential destination states. In this case the migration decision can be perfectly replicated.

Therefore, for each potential destination state j the predictive error $\upsilon_{i,j,\alpha_{i,t,T},t,T}$ can be formalized as the maximum of zero (no predictive error) or the difference of the respective preference value of state j and the respective preference value of the actual destination state d as follows

$$\upsilon_{i,j,t,T}\left(\alpha_{i,t,T}\right) = max\left\{0; \psi_{i,j,t,T}\left(\alpha_{i,t,T}\right) - \psi_{i,d_i,t,T}\left(\alpha_{i,t,T}\right)\right\} \tag{13}$$

where $\upsilon_{i,j,t,T}\left(\alpha_{i,t,T}\right)$ denotes the predictive error of the model as a function of the respective individual's risk-attitude $\alpha_{i,t,T}$ that relates to the migration decision of individual i, state j, concerning the migration decision taken at time t with planning period $(T - t)$.

While Equation (13) denotes the predictive error for one single state, the vector of predictive errors over all potential destination states reads as follows

$$\boldsymbol{\Upsilon}_{i,t,T}\left(\alpha_{i,t,T}\right) = \begin{pmatrix} \upsilon_{i,j=1,t,T}\left(\alpha_{i,t,T}\right) \\ \cdots \\ \upsilon_{i,j=50,t,T}\left(\alpha_{i,t,T}\right) \end{pmatrix}. \tag{14}$$

where $\boldsymbol{\Upsilon}_{i,t,T}\left(\alpha_{i,t,T}\right)$ denotes the vector of predictive errors as function of the respective individual's risk-attitude $\alpha_{i,t,T}$ that relates to the migration decision of individual i concerning the migration decision taken at time t with planning period $(T - t)$.

1.2.2 Number versus magnitude of predictive errors

In principle, predictive errors can be measured either by the pure number of errors (thereby ignoring their magnitude) or by weighting errors by their magnitude. In my analysis, I choose to weight errors by their magnitude because the pure number of errors might be misleading. To make this point clear, consider two scenarios resulting from two different parameter values on risk-attitude. In the first scenario, there are only a small number of predictive errors, but one error has a great magnitude. This means the model predicts the migrant to have a great preference for at least one other state over the actually chosen one. In the second scenario, a high number of predictive errors occur, but each is small in magnitude. Here the model predicts only a slight preference for other states than the one actually chosen.

The two scenarios show that the magnitude of the predictive errors cannot be neglected. Otherwise enormous predictive errors would be equally weighted like marginal small predictive errors. Since this makes no sense from an economic point of view, I decide to weight predictive errors by their magnitude in order to account for both criteria – number and magnitude.

1.2.3 Weighting of predictive errors: L_p-norm

Once I decided to weight predictive errors by their magnitude, they must be aggregated to a single error measure in order to be able to minimize predictive errors over all potential destination states.

Therefore, the L_p-norm which considers only the absolute value of errors and additionally enables the weighting of predictive errors v_{i,s,α_i} is applied in my analysis. The L_p-norm is defined as[236]

$$\left\|\boldsymbol{\Upsilon}_{i,t,T}(\alpha_{i,t,T})\right\|_p = \left(\sum_s \left|v_{i,j,t,T}(\alpha_{i,t,T})\right|^p\right)^{1/p}. \tag{15}$$

where p denotes a positive integer with $p = 1,2,\dots,\infty$.

In my study I apply the three most common L_p-norms as error measures, namely one-, two- and infinity-norms, i.e., $p = \{1,2,\infty\}$. The L_1-norm equally weights the absolute values of all predictive errors of all states. The L_2-norm puts greater weights on errors having a greater absolute value. It is also known as the standard Euclidean norm. Finally, the L_{max}-norm ($p = \infty$) puts all weight on the maximum absolute value of predictive error, i.e., the state with the greatest surplus in its preference value over the preference value of the actual destination state. The minimization of the L_{max}-norm can also be interpreted as minimizing the opportunity costs, i.e., the preference value of the next best foregone alternative compared to the actually chosen destination state.

Since there are no arguments why one of these L_p-norms would be preferable over the others, all three measures of predictive errors are equally considered in the reminder of my study.

1.3 Different ways to estimate risk-attitudes

The results of the empirical analysis crucially depend on the way risk-attitudes are estimated. The different ways to estimate risk-attitudes are due to (i) the different possibilities to specify the model, (ii) gain the data input, and (iii) the different estimation procedures available. Since it is not ex ante clear which way to estimate risk-attitudes is the right way, I decide to keep all possible variations in my study. This allows me to run a far ranging sensitivity analysis on all results of my study.

The different ways to estimate risk-attitudes relate of the following topics briefly summarized below (see Table 29, p. 150 and Table 30, p. 151).[237]

(i) Decision problems

Decisions based on personal versus family income: Migration decision based on Head's individual income versus family income (see Part A, Section 2.2.4).

---

[236] See Yang (2008), p. 29 and Wilhelm and Brüning (1992), p. 269.

[237] The different components its problems and solutions are discussed in detail in Part B of this study.

Three different planning periods: The theoretical migration decision model is applied to three different planning periods, namely one year, time until reaching full retirement age, and time until reaching life expectancy (see Part A, Section 2.2.2 and 2.4.2).

Two different risk-measures: The risk involved in the migration decision can be measured as variance of income or as semi-variance, i.e., the variance of short-fall below the expected income (Part A, Section Section 3.2).

(ii) Ways to gain data input

Six definitions of education: First, information about migrant's education is not available for every year but must be taken from earlier or later years of data. Second, several education variables exist in the data set that are not comparable to each other (see Table 14, p. 82 for a short overview and Part B, Section 2.3.3.6 for a detailed discussion). The six education definitions are applied only to migration decisions based on Head's individual income. For migration decisions based on family income, the analysis is restricted to the first education definition (Edu1). The reason for the restriction is that the robustness of the results concerning different education definitions can already be checked for migration decisions based on Head's income, and does not need to be checked again for migration decisions based on family income.

Weighted versus unweighted sample: Income parameters can be estimates based on weighted and unweighted samples of the American Community Survey (see Part B, Section 3.4.4.2).

Separate clustering for each year versus pooled clustering: To estimate income parameters for each combination of gender, age, and education from the American Community Survey, individuals in the sample must be clustered to groups for which income parameters are most similar. This can be done either separately for each year based on income parameters of that year or pooled over all years based on the mean income parameters over all years (see Part B, Section 3.4.4.2). Note that the four types of clustering discussed Part B, Section 3.4.4.2 already include the combination with weighted/unweighted income parameters.

Three time periods from which income parameters are estimated: Income parameters are estimated based (i) on annual income of the year of the move, (ii) based on three years of annual income data (namely, the actual year of the move, the year preceding, and the one following the move) without inflation adjustment, and (iii) based on three years of annual income data (namely, the actual year of the move, the year preceding, and the one following the move) with inflation adjustment. Inflation adjustment is performed using the Consumer Price Index (see Part B, Section 3.6).

(iii) Estimation procedures

3 ways to weight predictive errors: The most common L_p-norms are applied, namely the one-, two- and infinity-norms, i.e., $p = \{1,2,\infty\}$ (see Part C, Section 1.2.3).

Decision problems			**Ways to gain data input**				**Estimation procedure**
Income on which Head's migration decision is based	**Planning period**	**Risk-measure**	**Definition of education**	**Weighting of sample**	**Type of clustering**	**Time periods from which income parameters are estimated**	**Weighting of predictive errors**
Head's individual income (Ind_)	- One year, i.e., annual (Ann) - Time until reaching full retirement age, i.e., working life (Wor) - Expected remaining lifetime (Lif)	- Variance (Var) - Semi-variance, i.e., Lower Partial Moment 2 with reference value mean (LP2)	- Education 1 (Ed1) - Education 2 (Ed2) - Education 3 (Ed3) - Education 4 (Ed4) - Education 5 (Ed5) - Education 6 (Ed6)	- Weighted (Wei) - Unweighted (Unw)	- Separate clustering for each year (Sep) - Pooled clustering over mean of all years (Poo)	- Actual year of move, i.e., one year (One) - Actual, preceding and following year of move with inflation adjustment (Ad1) - Actual, preceding and following year of move without inflation adjustment (Ad2)	- L_1-norm (L1) - L_2-norm (L2) - L_∞-norm (Ma)
1	·3	·2	·6	·2	·2	·3	·3 = 1,296

Table 29: Overview of 1,296 ways to estimate risk-attitude for migration decisions based on Head's individual income.
Source: Own illustration. Codes in parentheses are abbreviations of each component that constitute the name of variables on risk-attitude. For example, all variables on risk-attitude starting with "Ind_" relate to decisions based on Head's individual income.

Decision problems			Ways to gain data input				Estimation procedure
Income on which Head's migration decision is based	**Planning period**	**Risk-measure**	**Definition of education**	**Weighting of sample**	**Type of clustering**	**Time periods from which income parameters are estimated**	**Weighting of predictive errors**
Family income (Fam_)	- One year, i.e., annual (Ann) - Time until reaching full retirement age, i.e., working life (Wor) - Expected remaining lifetime (Lif)	- Variance (Var) - Semi-variance, i.e., Lower Partial Moment 2 with reference value mean (LP2)	- Education 1 (Ed1)	- Weighted (Wei) - Unweighted (Unw)	- Separate clustering for each year (Sep) - Pooled clustering over mean of all years (Poo)	- Actual year of move, i.e., one year (One) - Actual, preceding and following year of move with inflation adjustment (Ad1) - Actual, preceding and following year of move without inflation adjustment (Ad2)	- L_1-norm (L1) - L_2-norm (L2) - L_∞-norm (Ma)
1	·3	·2	·1	·2	·2	·3	·3 = 216

Table 30: Overview of 216 ways to estimate risk-attitude for migration decisions based on family income.
Source: Own illustration. Codes in parentheses are abbreviations of each component that constitute the name of variables on risk-attitude. For example, all variables on risk-attitude starting with "Fam_" relate to decisions based on family income.

1.4 Numerical results of non-linear optimization

The Mathematics and Modeling software Maple 15 and its command NLPSolve are used to find the risk-attitude with an accuracy of thirty digits after the decimal point that relates to the smallest predictive error measure by three L_p-norms. The 1,512 ways to estimate risk-attitude result in a total of 481,680 parameter values on risk-attitude that have to be estimated (321 migration decisions based on Head's individual income estimated in 1,296 different ways (see Table 29, p. 150) and 304 migration decisions based on family income estimated in 216 different ways (see Table 30, p. 151)). Due to the high number of estimated parameter values, only a selection of representative examples is discussed in this section. For all 481,680 estimation results please refer to the data storage that comes along with this study. The corresponding files can be found in „\Alpha\(5) Alphas, full sample, all variables\Family\1_Family_Alphas collected from Maple.xlsx" and „\Alpha\(5) Alphas, full sample, all variables\Personal\1_Personal_Alphas collected from Maple.xlsx".

To get a first impression of the estimated risk-attitudes, Figure 20, p. 154 plots estimated risk-attitudes against their corresponding measurement of predictive errors exemplary for four ways to estimate risk-attitudes. It is noticeable, first, that all parameter values on risk-attitude are pretty small in absolute terms. More precisely, among all 1,512 ways to estimate risk-attitude, parameter values range from -0.002 to 0.004. The small values in absolute terms are not surprising but economically sensible for two reasons: First, recall that the preference value describes the trade-off of the conflicting goals expected value and variation of income, where risk-attitude is the preference weight put on the variation of income. Second, variation of income is measured by quadratic concepts variance or semi-variance, respectively, and is always much higher than the expected value. For example, annual income parameters in the data set amount on average to USD 36,252 for expected income in comparison to USD 1,734,334,982 for variance of income and USD 382,426,794 for semi-variance of income, respectively.[238] Consequently, the preference weight put on the variation of income must be much smaller than unity in order to gain some balance between the two conflicting goals expected value and variation of income.

Second, although small in absolute terms, all scatter plots in Figure 20, p. 154 reveal that the great majority of migrants has a positive parameter value for risk-attitude, i.e., they are risk-averse. This is

[238] Note that for longer planning periods the variation of income gets even greater in relation to expected income. Hence, parameters on risk-attitude get even smaller for longer planning periods compared to shorter planning periods. This can also be seen in Figure 20, p. 154 where Example 3 relates to the shortest planning period of one year (indicated by letters "Ann" in the variable's name) show the greatest parameter values on risk-attitude compared to Examples 1, 2 , and 4. Vice versa, Examples 1, 2, and 4 relate to a longer planning period of time until reaching full retirement age (i.e., working life indicated by letters "Wor" in the variable's name), and is associated with smaller parameter values of risk-attitude compared to Example 3.

true for all 1,512 ways to estimate risk-attitude. Depending on the way risk-attitude is estimated, the percentage of risk-averse migrants ranges from 80% to 98%.[239]

Third, migrants for which the migration decision can be replicated 100% can best be identified in Example 1. They are represented by the dots having a measured predictive error of zero. Although the number of perfectly predicted migration decisions is small no matter which of the 1,512 ways to estimate risk-attitude is considered, the effect is not easy to see in the other scatter plots. This is due to the concentration of dots for small positive values of risk-attitude.

[239] Although no literature on risk-attitude in the migration context exists, the finding that the great majority of people turns out to be risk-averse is in line with findings of empirical literature on risk-attitude in other non-migration domains (see, for example, Barsky, Juster, Kimball, and Shapiro (1997), Donkers, Melenberg, and van Soest (2001), Hartog, Ferrer-i-Carbonell, Jonker (2002), Ding, Hartog, Sun (2010), Holt and Laury (2002), Harrison, Lau, and Rutström (2007), and Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009)).

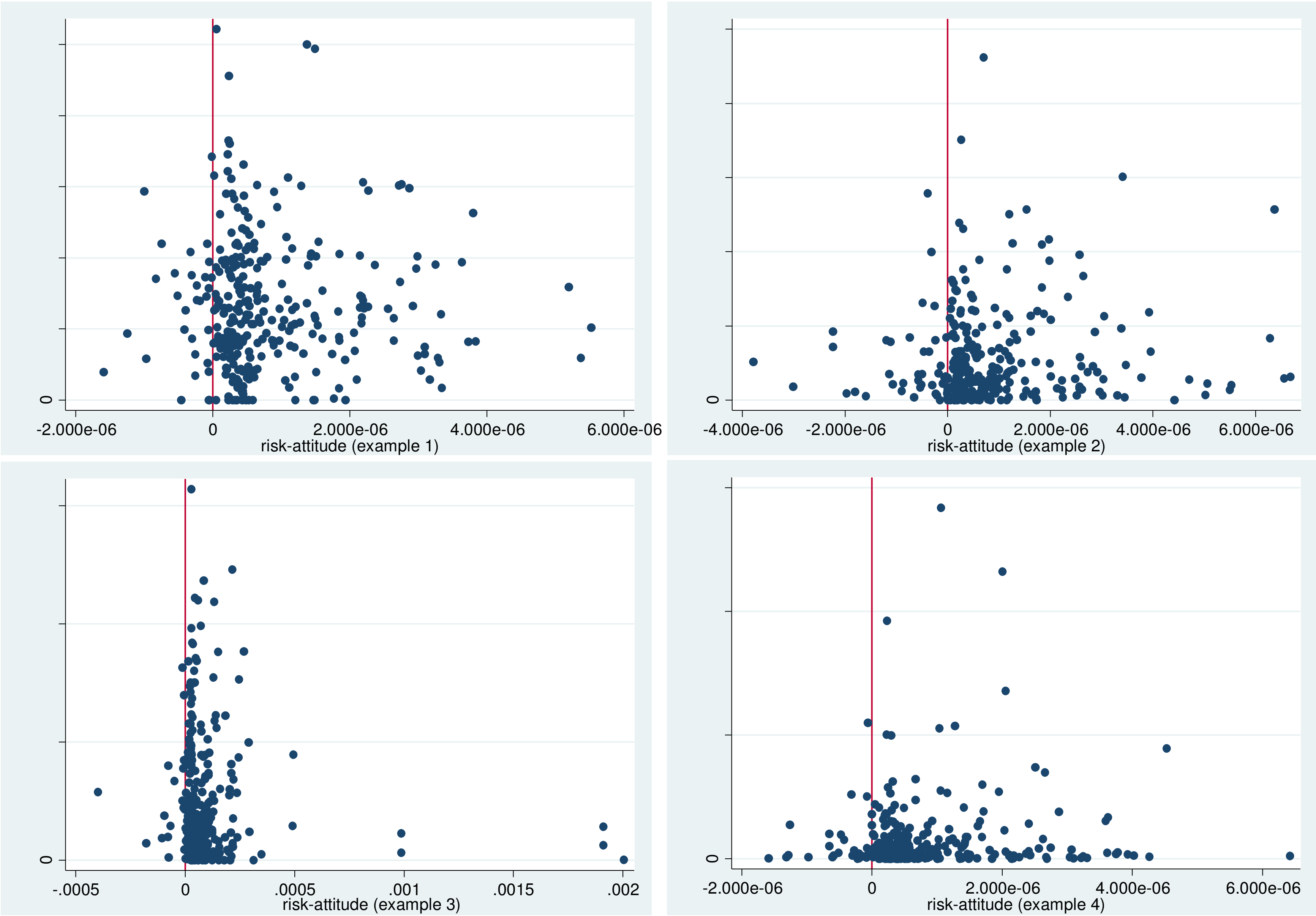


Figure 20: Scatter plots of risk-attitude and predictive errors for selected ways to measure risk-attitude.
Examples relate to variables of risk-attitude (1) Fam_WeiPooVarWorMaOneEd1, (2) Fam_WeiPooVarWorL1OneEd1, (3) Fam_WeiSepLP2AnnL1Ad2Ed1, and (4) Fam_WeiPooVarWorL2OneEd1, where the components of the variable's name are given in Table 29 and Table 30 on p. 150 f.

2 Preparing estimated risk-attitudes for further empirical analyses

2.1 Problems: outliers, point estimates, and optimal ranges

Problem 1: Outliers with respect to risk-attitude

The 1,512 ways to estimate risk-attitudes lead to some outliers with respect to risk-attitude. See for example, the lower left scatter plot of Figure 20, p. 154. Expressed in economic terms, these are the people that are extremely risk-averse or risk-seeking in comparison to the average migrant. Outliers raise three questions: Which criteria should be applied to identify outliers exactly? Do outliers bias the empirical analysis on the relation of socio-economic characteristics and risk-attitude? Should outliers be excluded from the analysis?

Problem 2: Mixture of single unique solutions and ranges of solutions for risk-attitudes

The non-linear optimization does not always result in a single unique solution but also in ranges of solutions. Ranges of solutions exist for migrants whose migration decision can be replicated 100%. Although the perfect replication of the migration decision is a wanted property, optimal ranges are a problem for further empirical analyses because no clear parameter value of risk-attitude can be assigned to these migrants.

Problem 3: Missing interval boundaries of ranges of solutions for risk-attitudes

Unfortunately, the exact interval boundaries for ranges of solutions are not reported by Maple. Although the missing interval boundaries can be determined in general, this is not feasible for the total of all 481,680[240] parameter values on risk-attitude due to computer capacity restraints. Therefore, I distinguish between two different types of problems relating to missing interval boundaries as follows:

Problem 3a: Ranges of solutions for risk-attitudes that do not transgress zero

Even if the exact interval boundaries are not known, ranges of optimal risk-attitudes that do not transgress zero can still be interpreted in economic terms since positive values of risk-attitude relate to risk-aversion and negative values in the same variable relate to risk-seeking.

[240] The 481,680 parameter values on risk-attitude result from 321 migration decisions based on Head's individual income estimated in 1,296 different ways (see Table 29, p. 150) and 304 migration decisions based on family income estimated in 216 different ways (see Table 30, p. 151).

Problem 3b: Ranges of solutions for risk-attitudes that transgress zero

Once optimal ranges of risk-attitudes transgress zero, it is not clear anymore whether the migrant is risk-averse (positive values), risk-seeking (negative values) or risk-neutral (value of zero).

2.2 Solution to Problem 3b: deleting ranges that transgress zero

Problems 1 and 2 in Part C, Section 2.1 can only be tackled once it is decided on how to deal with missing interval boundaries of the optimal ranges of solutions (Problems 3). I start with the more critical problem (Problem 3b) since all remaining problems cannot be solved before Problem 3b is solved.

The problem of optimal ranges only relies to migration decisions that can be replicated 100%. For them it is checked whether the corresponding ranges transgress zero. For example, if a predictive error of zero relates to an optimal risk-attitude of 0.00000135, it is checked whether the measured predictive error relating to a risk-attitude parameter of -0.00000001 is still zero. If this is the case, the respective migration decision includes an optimal range that transgresses zero. In abstract terms, the procedure to check whether a range transgresses zero can be described as follows: First, I switch to the opposite sign of zero compared to the risk-attitude value reported by Maple. Second, given the reversed sign, I chose the parameter value on risk-attitude that lies next to zero with an accuracy of 8 digits after the decimal point, i.e., a value of positive or negative 0.00000001. I arbitrarily choose an accuracy of 8 digits to capture ranges that only slightly cross zero. If lower accuracy was chosen, it could be that ranges that actually transgress zero but are not detected although they actually transgress zero. At the same time I did not choose a greater accuracy of 9 digits or more since most parameter values found by Maple are no closer to zero. Third, I calculate the measured predictive error corresponding to positive 0.00000001, or negative 0.00000001 respectively. If it equals zero, the range obviously transgresses zero.

Fortunately, depending on the way risk-attitude is estimated only 0.7% to 1.3% of all risk-attitudes possess optimal ranges that transgress zero. Due to this low number and because it cannot be differentiated between risk-averse, risk-neutral, and risk-seeking individuals, they are excluded from the analysis.

2.3 Solution to problems 1 to 3a: binary scaling of risk-attitudes

I decide to solve Problems 1 to 3a by transforming all remaining solutions on risk-attitudes (i.e., single unique solutions and ranges of solutions that do not transgress zero) to a binary scaled variable on risk-attitude applying a threshold value of zero for the following reasons:

First, starting from an economic point of view, the decisive threshold value for the estimated parameter on risk-attitude is zero because zero separates risk-averse (positive values) from risk-neutral (value of zero), and risk-seeking (negative values) individuals. Second, risk-neutral migrants cannot be found in the data set which means only two categories of people are left. Third, a binary coding makes it possible to incorporate outliers instead of deleting them without biasing the results. At the same time, the problems of how to define outliers is avoided (Problem 1). Fourth, binary coding with threshold value zero makes single unique solutions and ranges of solutions not transgressing zero comparable (Problem 2) irrespective of their exact interval boundaries. This means fifth, the exact interval boundaries of optimal ranges do not need to be known (Problem 3). The latter is especially important since ranges relate to migration decisions that can be replicated 100%. If these records were to be excluded, the most interesting cases would be missing. Sixth, even if interval boundaries were to be known, a more detailed classification of people into a multilevel ordinal scale does not make sense due to the small number of migrants that show risk-seeking behavior, i.e., have a negative parameter value of risk-attitude $\alpha_{i,t,T}$. This is true for all 1,512 ways to measure risk-attitudes.

The binary scaling with threshold value zero is performed by three different transformation rules: clear-cut at threshold value zero, deleting 5% weakest risk-attitudes, and keeping only the most extreme half of risk-averse and risk-seeking migrants.

First rule of transformation: Clear-cut at threshold value zero

Numerical results on risk-attitudes are transformed to a binary variable by dividing all people into risk-averse and risk-seeking migrants applying a threshold value of zero. The resulting binary variable will be the standard dependent variable in the following analysis.[241]

Second rule of transformation: deleting 5% weakest risk-attitude

I want to acknowledge that there are certain concerns in choosing the simple first transformation rule. This relates to the often very small risk-attitude parameters found by Maple (see Part C,

[241] In the data set the transformed binary variable is indicated by “2Kat3”.

Section 1.4) of which many are close to zero. For them it cannot be ruled out that applying a strict threshold value of zero leads to a false classification of individuals due to limited precision of input parameters and general limitations of the model. In order to avoid such a misclassification, individuals having parameter values on risk-attitude close to zero could be deleted to a certain degree. Since interval boundaries for certain individuals are still not known (Problem 3a from Part C, Section 2.1), this raises the question of which individuals exhibit the weakest degree of risk-attitude, i.e., whose parameter values on risk-attitude are closest to zero.

I solve this problem, by replacing all parameter values that (i) relate to optimal ranges and (ii) do not transgress zero (Problem 3a), by their interval boundary lying next to zero. This means, the recoded risk-attitude has the right sign (because all risk-attitude parameters in the range have the same sign), but potentially underestimates the degree of risk-attitude. For example, if the predictive error of a migration decision is minimized for all risk-attitudes ranging from 0.00000135 to 0.004, the migrant is risk-averse concerning his migration decision. The recoded risk-attitude is set to 0.00000135. This value also indicates risk-averse migration behavior, but the extent of risk-aversion might be underestimated because the same minimal predictive error is reached for higher degrees of risk-aversion relating to greater values of the risk-parameter. Note that the interval boundary is again determined with an accuracy of 8 digits after the decimal point for two reasons: First, an accuracy of 8 digits after the decimal point results in interval boundaries that are comparable to risk-attitude values found by Maple since the majority of values found by Maple has their first non-zero figure at the 6th digit after the decimal point. Second, a higher precision of the interval boundary is associated with considerably more computation time needed for the estimation procedure. An accuracy of 8 digits after the decimal points balances precision and computation time needed for the estimation.

Once the recoding of optimal ranges to interval boundaries lying next to zero has been performed, the second transformation rule picks up the concerns related to the first transformation rule as follows: First, the 5% of all migrants that have the smallest parameter value of risk-attitude in absolute terms (i.e., the 5% lying closest to zero) are deleted from the sample. Those are the 5% of all migrants that are closest to being risk-neutral, i.e. have the smallest degree of risk-attitude. The left over 95% are binary coded using the threshold value of zero.[242]

[242] In the data set the transformed binary variable is indicated by “2Kat1”.

<u>Third rule of transformation: Keeping only most extreme half of risk-averse and risk-seeking migrants</u>

Finally, the third transformation rule addresses the concerns of misclassification by restricting the analysis to those migrants that belong (i) either to the most extreme half among the risk-averse or (ii) the most extreme half among the risk-seeking migrants.[243] In other words, the sample is reduced to 50%. Since the sample is now restricted to migrants whose risk-attitude is more pronounced, the effect of socio-economic characteristics on risk-attitudes should also be more pronounced.

2.4 4,536 ways to estimate risk-attitude

The three transformation rules to gain a binary coded variable on risk-attitude are applied to all 1,512 different ways to estimates risk-attitudes as outlined in Table 29, p. 150 and Table 30, p. 150. Consequently, the analysis on the relation of socio-economic characteristics and risk-attitudes must be performed for 4,536 dependent variables all of which measure risk-attitude. Table 31, p. 160 summarizes the resulting combinations.

[243] In the data set the transformed binary variable is indicated by “4Kat2”.

Decision problems			**Ways to gain data input**				**Estimation procedure**	**Transformation to binary dependent variable**
Income decision is based on	**Planning period**	**Risk-measure**	**Definition of education**	**Weighting of sample**	**Type of clustering**	**Time periods from which income parameters are estimated**	**Weighting of predictive errors**	**Transformation rules**
Head's individual income (Ind_)	- One year, i.e., annual (Ann) - Time until reaching full retirement age, i.e., working life (Wor) - Expected remaining lifetime (Lif)	- Variance (Var) - Semi-variance, i.e., Lower Partial Moment 2 with reference value mean (LP2)	- Education 1 (Ed1) - Education 2 (Ed2) - Education 3 (Ed3) - Education 4 (Ed4) - Education 5 (Ed5) - Education 6 (Ed6)	- Weighted (Wei) - Unweighted (Unw)	- Separate clustering for each year (Sep) - Pooled clustering over mean of all years (Poo)	- Actual year of move, i.e., one year (One) - Actual, preceding and following year of move with inflation adjustment (Ad1) - Actual, preceding and following year of move without inflation adjustment (Ad2)	- L_1-norm (L1) - L_2-norm (L2) - L_∞-norm (Ma)	- Threshold value zero (_2Kat3) - Deleting 5% weakest risk-attitudes, threshold value zero (_2Kat1) - Threshold value zero, only most extreme half of risk-averse and risk-seeking migrants (_4Kat2)
Family income (Fam_)	- One year, i.e., annual (Ann) - Time until reaching full retirement age, i.e., working life (Wor) - Expected remaining lifetime (Lif)	- Variance (Var) - Semi-variance, i.e., Lower Partial Moment 2 with reference value mean (LP2)	- Education 1 (Ed1)	- Weighted (Wei) - Unweighted (Unw)	- Separate clustering for each year (Sep) - Pooled clustering over mean of all years (Poo)	- Actual year of move, i.e., one year (One) - Actual, preceding and following year of move with inflation adjustment (Ad1) - Actual, preceding and following year of move without inflation adjustment (Ad2)	- L_1-norm (L1) - L_2-norm (L2) - L_∞-norm (Ma)	- Threshold value zero (_2Kat3) - Deleting 5% weakest risk-attitudes, threshold value zero (_2Kat1) - Threshold value zero, only most extreme half of risk-averse and risk-seeking migrants (_4Kat2)

Table 31: Overview of 4,536 ways to estimate risk-attitude separate for migration decisions based on Head's individual income and family income.
Source: Own illustration. Codes in parentheses are abbreviations of each component that constitute the name of variables on risk-attitude. For example, all variables on risk-attitude starting with "Fam_" relate to decisions based on family income.

3 Is risk relevant for the migration decision?

Scatter plots in Part C, Section 1.4 have shown that risk-attitudes are (i) small in magnitude but seem to be different from zero, and (ii) seem to be located primarily in the positive quadrant. Hence, this section is concerned with the statistical test of this presumption, i.e., gives an answer to the question whether risk actually plays a role in the migration decision (second research question of my study).

3.1 Hypotheses

In order to elaborate whether risk actually plays a role in the migration decision, it must statistically be tested whether migrants are on average different from being risk-neutral. This null hypothesis can be tested with two types of data: first, with the original results on risk-attitude reported by Maple, or second, with the transformed binary coded variables. The first type of data is afflicted with the problem on how to deal with outliers and how to handle optimal ranges that do not include zero. The second type of data means a loss of information because it treats all risk-averse (risk-seeking) migrants equally although there are huge differences in the degree of risk-aversion (risk-seeking). Since both approaches have their pros and cons, I decide to test the null hypothesis in two different ways.

Concerning the first type of data, results on risk-attitudes initially reported by Maple are cleaned by (i) deleting optimal ranges that include zero, and (ii) replacing risk-attitude of other optimal ranges by their upper/lower interval boundary lying next to zero as described in Part C, Sections 2.2 and 0. Since the cleaned data is scaled metrically, a simple t-test is run on each of the 1,512 ways to estimate risk-attitude. The corresponding null hypothesis states

$H0_{t\text{-}test}$: The average risk-attitude of migrants is zero.

In the second type of data, risk-attitudes are scaled binary as described in Part C, Section 2.2. For each of the three transformation rules the following null hypothesis is tested separately using a binomial test as follows

$H0_{Binomial}$: Risk-averse and risk-seeking migrants are equally likely to occur.

3.2 Results

3.2.1 Risk matters

The results on significance of risk in the context of economic are striking. The $H0_{Binomial}$ can be rejected at a 1% significance level for all 4,536 ways to estimate risk-attitude by a binary variable. Furthermore, $H0_{t\text{-}test}$ can be rejected at 1% level of significance for all 1,296 migration decisions based on Head's individual income and for 97% of all 216 migration decisions based on family income. More precisely, there are only 6 ways to estimate risk-attitude for which the 1% significance level does not hold; only 5 of them do not reach a significance level of 10% (see Table 32, p. 163).

To see more details on the test statistics, investigate exemplary for the 4,536 ways to estimate risk-attitude, the four ways to estimate risk-attitude that have already been illustrated by the scatter plots in Figure 20, p. 154. For them Table 32, p. 163 lists the corresponding parameter values of risk-attitude and predictive errors in a stylized fashion. The bottom row of Table 32, p. 163 reveals that for these 4 ways to estimate risk-attitude the null hypothesis using both data sets and its corresponding test statistics can be rejected at a 1% significance level.[244]

Those few exceptions for which risk-attitudes are not found to be significantly different from zero do not follow any systematic pattern (see Part C, Section 3.2.2 for a detailed discussion). Hence, concerning the second research question of my study, I conclude that risk actually plays a significant role in the migration decision.

Since for all ways to estimate risk-attitudes that are significantly different from zero (i) more than half the migrants are risk-averse, and (ii) mean risk-attitudes are positive, I can even make a stronger statement: Migrants are not only significantly different from being risk-neutral, but exhibit significant risk-aversion on a 1% significance level in the migration context.[245]

[244] The number of observations for which risk-attitudes are reported in Table 32, p. 163 is always smaller than 304 because optimal ranges of risk-attitudes that include zero have been deleted from the sample (see, for example, observation 278). In contrast, for observation 275 in Table 32 the optimal range of risk-attitude does not include zero. Therefore, the lower boundary of the optimal range is reported as risk-attitude.

[245] Note that an additional hypothesis test on the corresponding one-sided null hypothesis (null hypothesis: Migrants are on average risk-seeking) is not necessary since the only difference between the two-sided and one-sided test is that the resulting p-values for the one-sided test amount to exactly half of the two-sided test.

	Example 1		Example 2		Example 3		Example 4	
observation	**risk-attitude**	**predictive error**	**risk-attitude**	**predictive error**	**risk-attitude**	**predictive error**	**risk-attitude**	**predictive error**
1	0.00000067	161363	0.00000036	230147	0.00001350	34080	0.00000039	35309231832
2	-0.00000005	39793	-0.00000065	72367	0.00006958	19254	-0.00000019	3182644116
3	0.00000045	86607	0.00000067	323227	0.00004512	28271	0.00000057	21165269041
…	…	…	…	…	…	…	…	…
276	0.00000216	61018	0.00000377	99885	0.00020384	5051	0.00000307	5327448230
275	0.00000023	0	0.00000023	0	0.00004144	6215	0.00000023	0
276	0.00000216	61018	0.00000377	99885	0.00020384	5051	0.00000307	5327448230
277	0.00000035	59479	0.00000023	449636	0.00007743	21278	0.00000037	16347477451
278	n.a.	0	n.a.	0	n.a.	0	n.a.	0
279	0.00000220	119341	0.00000554	1310483	0.00098773	3451	0.00000378	87815436281
280	0.00000070	131484	0.00000041	1149079	0.00005078	37316	0.00000037	90391218218
281	0.00000014	222705	0.00000001	2784551	0.00003159	99703	0.00000011	382671939517
282	0.00000153	102719	0.00000027	219441	0.00004503	26153	0.00000076	23828715396
…	…	…	…	…	…	…	…	…
300	0.00000024	146619	0.00000015	417673	-0.00000246	14381	0.00000017	46232058051
301	0.00000048	365146	0.00000043	4779621	0.00005334	160226	0.00000041	1005026558200
302	0.00000299	79591	0.00000006	605879	0.00002562	27485	0.00000092	37422470867
303	0.00000208	148564	0.00000040	390904	0.00002773	14303	0.00000085	43330714371
304	0.00000060	148854	0.00000025	1724678	0.00002728	58285	0.00000041	170957674494
number of risk-attitudes	302		302		300		302	
number of risk-seekers	32		41		21		28	
mean	0.00000083		0.00000082		0.00009348		0.00000081	
standard deviation	0.00000105		0.00000137		0.00021662		0.00000103	
p-value $H0_{t\text{-}test}$	0.00		0.00		0.00		0.00	
p-value $H0_{Bintomial}$	0.00		0.00		0.00		0.00	

Table 32: Exemplary results of estimated risk-attitudes that are significantly different from risk-neutrality.
Source: Own estimations. Examples are the same like in Figure 20, p. 154. Example 1 relates to Fam_WeiPooVarWorMaOneEd1, example 2 to Fam_WeiPooVarWorL1OneEd1, example 3 to Fam_WeiSepLP2AnnL1Ad2Ed1, example 4 to Fam_WeiPooVarWorL2OneEd1. Variable names are read as follows: Fam_ denotes decisions based on family income, Wei estimations based on weighted samples, Poo (Sep) pooled clustering over all years (separate for each year), Var (LP2) risk-measure variance (lower partial moment 2), Wor (Ann) planning period of working life (annual), Ma (L1, L2) respective L_p-norm, One (Ad2) income data of one year (three years inflation adjusted), Ed1 education definition 1 (see Table 29 and Table 30, p. 150 for an overview).

3.2.2 Unsystematic exceptions where risk does not matter

Among the 6,048 statistical tests run (4,536 binomial tests and 1,296 plus 216 t-tests) there are only six ways to estimate risk-attitudes for which risk-attitudes are not significantly different from risk-neutrality at significance level 1% (see Table 33, p. 166 at the end of this section). Still, if the related six ways to estimate risk-attitudes followed any systematic in the sense that, for example, risk played no role when measured by variance, the question on whether risk played a role in the migration decision could not be answered per se but would depend on the way risk-attitude is estimated. Put differently, a clear-cut statement regarding risk matters irrespective of the way risk-attitudes are estimated is impossible in this case. Therefore, a systematic pattern among the non-significant ways to estimate risk-attitude must be ruled out.

Similarities and differences of the six exceptions

The six ways to estimate risk-attitude that do not result in risk-attitudes significantly different from risk-neutrality as tested by null hypothesis $H0_{t\text{-}test}$ are listed in Table 33, p. 166. All six ways relate to migration decisions based on family income (indicated by "Fam_"), income parameters estimated based on a weighted samples (indicated bei "Wei") and a clustering separate for each year (indicated by "Sep"), risk-measure variance (indicated by "Var"), a planning period until reaching full retirement age (indicated by "Wor"), and "Ed1" since for migration decisions are based on family income only education definition one is applied.

Besides this similarity (indicated by Fam_WeiSepVarWor), the six ways to estimate risk-attitude listed in Table 33, p. 166 differ in the way predictive errors are measured (L_{max}-norm indicated by "Ma", L_1-norm indicated by "L1", and L_2-norm indicated by "L2") and whether income parameters are estimated based on income data of one single year (indicated by "One") or income data over three years with inflation adjustment (indicated by "Ad1").

Exceptions are due to an unlucky constellations of single components

If risk was not relevant due to one single component, like variance, half of the 1,512 ways to estimate risk-attitude would have been insignificant rather than six of them. Obviously, it is not one single component like variance that results in risk-attitudes that are not significantly different from risk-neutrality.

The same argument holds for two components. For example, if risk was not relevant for the combination of variance and a planning period of time until reaching full retirement age (working life), one quarter of the 1,512 ways to estimate risk-attitude would have been insignificant rather than only six of them. This is not the case here. Obviously, it is not the combination of two components that results in risk-attitudes that are not significantly different from risk-neutrality.

The same argument holds for combinations of up to 6 components. Only if 7 components (excluding the 3 ways predictive errors are estimated by L_p-norms) out of 8 possible components are combined in a very specific manner, insignificance occurs as follows:

The first combination of 7 components that result in insignificant results irrespective the 8th components (L_p-norm) is the combination of risk-attitudes estimated based on family income (Fam), income parameters estimated based on a weighted samples (Wei) and a clustering separate for each year (Sep), risk-measure variance (Var), a planning period until reaching full retirement age (Wor), income parameters estimated from three years of income data adjusted by inflation (Ad1), and education definition one (Ed1). This combination relates to the first three ways to estimate risk-attitudes listed in Table 33, p. 166.

The second combination of 7 components that result in insignificant results irrespective the 8th components (L_p-norm) is the combination of risk-attitudes estimated based on family income (Fam), income parameters estimated based on a weighted samples (Wei) and a clustering separate for each year (Sep), risk-measure variance (Var), a planning period until reaching full retirement age (Wor), income parameters estimated from one year of income data (One), and education definition one (Ed1). This combination relates to the last three ways to estimate risk-attitudes listed in Table 33, p. 166.

Since these exceptions relate to (i) a combination of 7 out of 8 possible components, and (ii) only 2 out of 504 possible ways[246] to combine these 7 components, I consider the exceptions from the significant results presented in Part C, Section 3.2.1 to be due to an unlucky constellation of components and not being due a systematic effect.

[246] Of the all together 8 components as listed in Table 29, p. 150 and Table 30, p. 151 only 1 component, namely the way predictive errors are weighted, is not fixed for the 6 exceptions. The remaining 7 components can be combined in 504 ways, i.e., 1,512 total combinations divided by the 3 ways predictive errors can be estimated.

Way to estimate risk-attitude	n	number risk-averse	number risk-seeking	mean risk-attitude	standard deviation of risk-attitude	p-value $H0_{t\text{-}test}$	p-value $H0_{Binoml}$
Fam_WeiSepVarWorMaAd1Ed1	300	279	25	0.0000006	0.00000587	0.07	0.00
Fam_WeiSepVarWorL2Ad1Ed1	300	271	33	0.0000005	0.00000584	0.15	0.00
Fam_WeiSepVarWorL1Ad1Ed1	300	256	48	0.0000005	0.00000592	0.18	0.00
Fam_WeiSepVarWorMaOneEd1	301	276	28	0.0000003	0.00000997	0.61	0.00
Fam_WeiSepVarWorL1OneEd1	301	267	37	0.0000003	0.00001001	0.64	0.00
Fam_WeiSepVarWorL2OneEd1	301	277	27	0.0000003	0.00000996	0.67	0.00

Table 33: Ways to estimate risk-attitudes for which risk-attitudes are not significantly different from risk-neutrality. Source: Own illustration based on own calculations. For an overview on how to read variables names see Table 29, p. 150 and Table 30, p. 150). N denotes the number of risk-attitudes that are considered for the statistical tests.

4 Risk-attitudes and socio-economic characteristics

Since the previous section has shown that risk actually plays a significant role in the migration decision, this section is concerned with the analysis of the relation of risk-attitudes and socio-economic characteristics of economic migrants (third research question of my study). Relating risk-attitudes and socio-economic characteristics is important because it contributes to one of the crucial questions in migration policy: How can the desired group of young and well-educated migrants be attracted?

To answer this question, Part C, Section 4.1 discusses which variables on socio-economic characteristics are included in the study. Part C, Section 4.2 explains the methodology of the extreme bounds analysis, and Part C, Section 4.3 presents empirical results on the relation of socio-economic characteristics and risk-attitudes in the context of economic migration.

4.1 Overview of 32 independent variables

The theoretical migration decision model of my analysis does not specify the socio-economic variables that relate to risk-attitudes. Therefore, a wide variety of socio-economic characteristics that were found to relate to risk-attitude in the empirical literature is entered as main effects. In addition to findings of the previous literature, several two-way interactions are also included. For all 32 variables entered in my study this section presents previous findings of the literature including my expectation concerning sign of the corresponding coefficient.

4.1.1 Main effects

The three most common socio-economic characteristics in the literature on risk-attitude are gender, education, and age. They are found to be significantly related to the willingness to take risks in general (i.e., not specific-domain) or in other non-migration domains by almost all studies. Therefore, variables on gender, age, and education will be discussed first. After that several other socio-economic characteristic that possibly relate to risk-attitudes in the migration context are examined. For an overview of all variables entered as main effects see Table 34, p. 174 at the end of this section.

Male

The most common explanatory variable on risk-attitude is gender. Literature widely agrees that women are significantly more risk-averse than men in all domains ever surveyed.[247] Although it is often hypothesized that this results from an omitted variables problem, no such findings have been reported in the literature so far.[248] This may be due to the mostly small number of explanatory variables combined with mostly missing interaction terms. A shortcoming that is tackled in this study by accounting for altogether 32 explanatory variables including a variety of interaction terms.

Concerning the domain of migration, I do not believe that men and women behave differently in taking the migration decision if it is controlled for education, age, and family ties. Still, in order to check the results found for other domains, a dummy variable on male Heads, namely Male, is entered in the analysis without a predetermined expectation.

Education

Although commonly surveyed, the effect of education on risk-attitudes found in the literature is not consistent. Only a slight majority of papers discover a significant negative effect of education on risk-aversion[249], while sometimes no effect is found at all.[250]

In the migration context, I find it reasonable to believe that education is related to risk-attitudes. As people acquire higher education levels they might be more capable to realize and evaluate the risk involved in the migration decision. This might give them the feeling of controlling the risk resulting in a higher willingness to take risks. Quite the opposite might hold for less educated people. They might see the risk but are not able to control it and, therefore, prefer not to take the risk. Accordingly, education is entered as independent variable in this analysis with the expectation that education is negatively related to risk-aversion.

---

[247] See for example, Barsky, Juster, Kimball, and Shapiro (1997), Powell and Ansic (1997), Donkers, Melenberg, and van Soest (2001), Hartog, Ferrer-i-Carbonell, and Jonker (2002), Weber, Blais, and Betz (2002), Johnson, Wilke, and Weber (2004), Wik, Kebede, Bergland, and Holden (2004), Nicholson, Soane, Fenton-O'Creevy, and Willman (2005), Bonin, Constant, Tatsiramos, and Zimmermann (2006), Grazier and Sloane (2006), Jaeger, Bonin, Dohmen, Falk, Huffman, and Sunde (2007), Jaeger, Dohmen, Falk, Huffman, Sunde, and Bonin (2008), Badunenko, Barasinska, and Schäfer (2009), Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009), Umblijs (2012). An exception who did not find a significant relation is Harrison, Lau, and Rutström (2007).
[248] See for example, Badunenko, Barasinska, and Schäfer (2009), p.9.
[249] A negative effect of education on risk-aversion, i.e., a higher willingness to take risks for those with higher levels of education, is found, for example, by Donkers, Melenberg, and van Soest (2001), Bonin, Constant, Tatsiramos, Zimmermann (2006), Jaeger, Bonin, Dohmen, Falk, Huffman, and Sunde (2007), Jaeger, Dohmen, Falk, Huffman, Sunde, and Bonin (2008), Umblijs (2012). The opposite is found by Harrison, Lau, and Rutström (2007) while Ferrer-i-Carbonell, and Jonker (2002) do not find any significant effect.
[250] See for example, Wik, Kebede, Bergland, and Holden (2004).

The quantification of education used in this study refers to the highest level of education achieved by Head at the time of the move. It was originally scaled ordinally on an eight point scale (see Part B, Section 2.3.3.3.1). In order to include the education variable in the regression analysis, a dummy coding with seven dummies is necessary. This results in very low frequencies for some levels of risk-attitude which means estimated coefficients are not very reliable – sometimes coefficients cannot be estimated at all. The problem is even more severe when possible interaction effects are entered later on. Therefore, the original eight education levels are reduced to a three-point scale as follows[251]

- Less than high school graduate
- High school graduate/Associate's degree
- Bachelor's degree and higher.

For the dummy coding education level "Less than high school" is chosen as reference category. Dummy variable "Edu2" refers to a maximum education of high school graduation or having an associate's degree. Dummy variable "Edu3" refers to those having a college degree (Bachelor and higher).

Age

Age in its linear form has been found to be related significantly positive to risk-aversion by almost all studies.[252] For example, Donkers, Melenberg, and van Soest (2001), Nicholson, Soane, Fenton-O'Creevy, and Willman (2005), Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009), and Umblijs (2012) who survey risk-aversion in several other non-migration domains and general risk-aversion (i.e., risk-aversion that is no related to any specific domain).

To test previous findings for the migration context, age in its linear form is included as independent variable without any predetermined expectation.

AgeSquare

Recent studies have shown that a nonlinear effect of age on risk-attitudes in non-migration domains exists. Harrison, Lau, and Rutström (2007), for example, analyze age groups and find people in their

---

[251] Note that the education definition in Part B was needed to estimate income parameters as data input for the empirical analysis. Once data input is gained, the education definition of Part B is no longer required. In contrast, the education definition applied in Part C has another meaning. It is applied to migrants in order to analyze the relation of education and risk-attitudes. The necessity to reduce the levels of the education variables in Part C is due to the methodology used here which is a consequence of the numerical results. It is therefore discussed in this section rather than in Part B.

[252] A rare exception is Hartog, Ferrer-i-Carbonell, and Jonker (2002) who find a significant though unequivocal effect depending on the sample used or Wik, Kebede, Bergland, and Holden (2004) who do not find any linear effect.

middle-ages (40 to 50 years old) to be most risk-averse compared to younger and older people. Many other studies have included age in the form of a quadratic age term in order to capture the possibly parabolic type of relation between age and risk-attitudes. They find age squared to be a significant determinant but their results are inconsistent. Hallahan, Faff, and McKenzie (2004) conclude that for the domain of finance risk-aversion increases at an increasing rate as age increases.[253] Others like Bonin, Constant, Tatsiramos, and Zimmermann (2006) find similar effects for the domains of finance, car driving, and career for people in their early twenties and older.[254] In contrast, the same authors find opposing effects for the willingness to take risks in general (i.e. not domain-specific), in sports and leisure, and in trusting strangers. In these domains risk-aversion increases up to ages 27, 40, and 28, respectively and decreases at an increasing rate as age increases afterwards.[255]

The findings of the literature show, that first, besides the simple linear implementation of age at least its marginal effect must be included using age squared. Second, the effect of age on risk-aversion is domain-specific. Hence, results found in non-migration domains cannot to be transferred to the migration context and AgeSquared is entered in my study as independent variable without a predetermined expectation.

Family size: Number of family members in the wave before/after the move

The effect of family size on risk-attitudes has been surveyed mostly for farmers in less developed countries where Heads of bigger families showed significantly lower risk-aversion compared to Heads of smaller families.[256] The opposite effect is reported for studies in developed countries like Hallahan, Faff, and McKenzie (2004), Muñoz and González (2011), Xiao, Alhabeeb, Hong, and Haynes (2001), while Harrison, Lau, and Rutström (2007) and Säve-Söderbergh (2012) do not find any effect.

Concerning migration risk, I can imagine two opposing effects of family size. On the one hand, a higher number of family members might indicate a stronger family background where people support each other therefore reducing risk-aversion. On the other hand, more people could also mean more financial responsibility for Head which would increase risk-aversion. Because of the

---

[253] This holds for individuals being older than 4.6 years. I calculate this turning point by means of the regression coefficients estimated by Hallahan, Faff, and McKenzie (2004), p. 67.
[254] The willingness to take risks in the domains of car driving, finance investment, and career increases up to ages 20, 23, and 19, respectively. Afterwards the willingness to take risks in these domains decreases by an increasing rate as age increases. Turning points can be calculated by regression coefficients reported by Bonin, Constant, Tatsiramos, and Zimmermann (2006), p. 8. A similar relation was found by Säve-Söderbergh (2012) with a turning point at age 31.
[255] I calculate these turning points from regression coefficients reported by Bonin, Constant, Tatsiramos, and Zimmermann (2006), p. 8.
[256] See for example, Wik, Kebede, Bergland, and Holden (2004).

opposing arguments and the inconclusive findings in the literature the variable is added without any predetermined expectation.

Note that a variable for family size at the time of the move is not available since migration takes place between two waves of data of the Panel Study of Income Dynamics. Therefore, it is not clear whether the number of family members in the wave before or after the move is decisive for the migration decision. For that reason, both variables are added. This is not problematic as it can be hypothesized that both variables measure the same aspect: family size.

Single, Pair, and Family

Although family size has often been found to be significantly related to risk-attitudes, it might not be appropriate in the context of economic migration. Instead of family size itself, more precise indicators might be (i) the number of people actually moving with Head and (ii) their personal relation to Head. Both aspects are captured by dummy the variables Single, Pair, and Family that are included in my study for the following reason:

Head as decision-maker is responsible for (i) the migration decision and (ii) people directly affected by it. Certainly, those people moving together with Head as a family are directly affected by the migration decision. Hence, Head is responsible for them in taking the migration decision. For non-moving family members Head's responsibility is not clear. Consider the following two examples. First, it could be that non-moving family members are roommates without any closer relation to Head. Naturally, unrelated roommates are not affected by the Head's migration decision. Second, non-moving family members include those that decided to leave the family or refused to move. Here again, these individuals are not directly affected by the migration decision. In both examples, non-moving family members are not affected by the migration decision which means, concerning the migration decision Head is not responsible for them.

A single-move (dummy variable **Single**) takes place when Head moves on his own without members of the old or new family. In this case Head is only responsible for himself. A pair-move (dummy variable **Pair**) takes place when two people who are indicated to be a pair move together with or without newborns. Note that in contrast to other studies which included a dummy on marriage,[257] the pair dummy used in this study also includes unmarried partnerships. This is done with intention.

[257] See for example, Hallahan, Faff, and McKenzie (2004), Bonin, Constant, Tatsiramos, and Zimmermann (2006), Grazier and Sloane (2006), Harrison, Lau, and Rutström (2007), Bertocchi, Brunetti, and Torricelli (2008), Jaeger, Dohmen, Falk, Huffmann, Sunde, and Bonin (2008), Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009), Muñoz and González (2011), O'Donnell (2011), Buurman, Delfgaauw, Dur, and van den Bossche (2012).

The reason is that for the investigation period 2000 to 2009 it is questionable whether the migration decisions of Heads living in a partnership where people (i) live together, (ii) move together, and (iii) report their relation in a national survey, are significantly different from migration decisions taken by married Heads. Maybe this unreasonable distinction made in the previous literature is the reason for the ambiguous results found: Bonin, Constant, Tatsiramos, and Zimmermann (2006) and others[258] using data from the German Socio-Economic Panel (SOEP) wave 2004 found married individuals to be significantly more risk-averse than unmarried individuals in general, and in the domains driving, finance, sports and leisure, career, health, and trusting strangers. The complete opposite is reported for financial risk-taking of married individuals in Italy, Spain, Ireland, and the United Kingdom[259], while some authors do not find any significant relation of marriage and risk-attitude like Harrison, Lau, and Rutström (2007) or Buurman, Delfgaauw, Dur, and van den Bossche (2012).

Finally, family-moves (indicated by the dummy variable **Family**) take place when (i) either two people move together as a family who are no pair before or after the move, or (ii) more than two people move together as a family where two of them could be a pair but not necessarily have to. Family-moves are distinguished from pair-moves because it seems reasonable that it makes a difference whether migration decisions are taken on behalf of "only" the partner (and potentially newborns around the time of the move) or on behalf of more people for several reasons: First, by definition, Head is the one bearing the most financial responsibility for the family. This means, the greater the family, the higher the potential financial obligations of Head. This argument goes in the same direction like the argument on the number of dependents (usually children) in the household which is sometimes found in the literature, but again with ambiguous results. Some authors find the number of dependents significantly positive related to risk-aversion like Hallahan, Faff, and McKenzie (2004), Bonin, Constant, Tatsiramos, and Zimmermann (2006), or Buurman, Delfgaauw, Dur, and van den Bossche (2012), while others find the opposite effect[260] or no significant effect[261]. Another reason to distinguish between pair-moves and family-moves is that second, it is reasonable to assume that the sheer number of people moving with Head influences his migration decision for organizational reasons. For example, the more people moving together the more housing space is

---

[258] See Jaeger, Dohmen, Falk, Huffmann, Sunde, and Bonin (2008) and Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009) whose results are based on the same data set. The results were also confirmed by later studies on financial risk tolerance among Australians by Hallahan, Faff, and McKenzie (2004) and taking risky jobs by Grazier and Sloane (2006).

[259] For Italy see Bertocchi, Brunetti, and Torricelli (2008), for Spain Muñoz and González (2011), and for Ireland and the United Kingdom O'Donnell (2011).

[260] An example is Bertocchi, Brunetti, and Torricelli (2008).

[261] An example is Säve-Söderbergh (2012).

needed in the destination area, the more things have to be moved, the more interests have to be reconciled etc.

Concerning the sign of variables Single, Pair, and Family there are principally two opposing arguments: On the one hand, a higher level of responsibility due to a higher number of people moving together suggests that single-movers exhibit a smaller risk-aversion than all family- and pair-movers, while family-movers exhibit a greater risk-aversion than all other Heads. This argument has been pronounced so far. On the other hand, moving together with more people could also mean that family members support each other. This would suggest that single-movers who are on their own exhibit greater risk-aversion compared to all others. Consequently, all three variables are entered into the analysis without a predetermined expectation.

Divorce

Finally, a dummy on Heads that separated or get divorced from their former partner between the wave before and after the move (dummy variable Divorce) is introduced as last main effect. To the best of my knowledge separation and its relation to any kind of risk-attitude has not been surveyed so far. Still, the variable is included in this study because it seems reasonable to me that separation could well interfere with economic reasons to move. Note, that Divorce is not restricted to legal divorce but to separation from married and non-married partners. This is in line with the definition of pair-moves that is also not restricted to married partnerships.

Concerning the relation of Divorce and risk-attitude I can think of several arguments. People who just separated from their old partner could be willing to take any risk just to leave the old partner. This could either result in especially risky decisions or in decisions taken without considering risk at all. Also, the experience of separation and its psychological burden could change people’s character in the sense that they are not able to take any risk any more. Although both arguments result in different expectations concerning the sign of the coefficients, they both highlight the difference of divorced/separated Heads compared to others. Therefore, the Divorce dummy is entered without any predetermined expectation.

Summary on main effects

Table 34, p. 174 gives an overview of all variables entered as main effect.

Variable name	Explanation
Male	Dummy on male Heads
Edu2	Dummy on having a “High school diploma or an associate’s degree”; reference category “Less than high school”
Edu3	Dummy on having a “bachelor’s degree or higher”; reference category “Less than high school”
Age	Age of Head at the time of the move
AgeSquared	Age squared
Number of family members in wave before the move	Number of family members in the family Head was associated with in the wave before the move (first proxy for family size)
Number of family members in wave after the move	Number of family members in the family Head was associated with in the wave after the move (second proxy for family size)
Single	Dummy on Heads moving on their own (single-move)
Pair	Dummy on Heads moving together with their partner (pair-move)
Family	Dummy on Heads moving with their family (family-move)
Divorce	Dummy on Heads who separated from their old partner between the wave before and after the move

Table 34: Main effects entered as independent variables.

4.1.2 Two-way interaction effects

Although interaction effects are rare in the empirical literature on risk-attitude, it seems reasonable in general to believe that the effects of many variables depend on the value of yet another variable, i.e., interaction effects should be included in the analysis. Speaking technically, an interaction effect exists if the effect of one variable (e.g., education) is estimated by separate regressions for each parameter value of another variable (e.g., gender), and the resulting regression lines are not parallel to each other.[262] Note that including an interaction effect in the regression controls for possible interactions, but if the interaction effect does not exist, it does not bias results of other regression coefficients.

Consequently, a wide variety of two-way interactions effects is included in the analysis, but I do not combine all main effects discussed above for two reasons: First, the combination of dummy variables resulting from the same categorical variable does not make sense because these interactions will

[262] See, e.g., Lomax and Hahs-Vaughn (2012), p. 89.

always equal zero, e.g., SingleFamily. In my study, this refers to categorical variables education and its three dummies as well as the Move Context Variable with its Dummies Single, Pair, and Family (see Part B, Section 2.4). Second, only those interaction effects that are worth investigating from an economic point of view are included. Therefore, those interactions included in the analysis and their economic interpretations are discussed below.

Except for the two variables on family size before and after the move and the variable on age square all main effects discussed in the previous section are entered in all possible combinations of two. I decide to exclude the variables on family size before and after the move because dummy variables Single, Pair, and Family principally measure the same effect, and are more precise at the same time. Recall, that they only consider family members that are directly affected by the migration decision, and additionally capture personal relations with Head. Furthermore, the variable AgeSquare is excluded from building two-way interactions because the economic insight of such a two-way interaction is very limited without the corresponding two-way interaction with age.

4.1.2.1 Interactions with Male

Interaction effects MaleSingle, MalePair, MaleFamily

Very few papers have surveyed whether gender differences are due to other explanatory variables on risk-attitudes. In other word, only very few studies include two-way interactions with Male. With one exception, those studies available survey financial risk-taking and are concerned with the interaction of gender and marriage. The first to report such an interaction effect were Sundén and Surette (1998). Later Barber and Odean (2001) found gender differences in risk-taking to be greater for married compared to unmarried individuals. The opposite is reported by Säve-Söderbergh (2012) who concludes that marriage makes men more risk-averse, while women become less risk-averse. No difference between married and unmarried individuals could be found by Yao and Hanna (2005). Dividing all respondents into four groups depending on marital status and gender, they found unmarried men to be least risk-averse, followed by married men and unmarried women with married women being the most risk-averse group.

Obviously no clear conclusion on the interaction of marriage and gender on risk-attitudes can be drawn from the existing literature, but the findings suggest that the interaction is worth to be investigated. Therefore, the corresponding interactions that are included without predetermined expectation are MaleSingle, MalePair, and MaleFamily.

Interaction effect MaleDivorce

If different gender effects for married and unmarried individuals are found in the literature, it seems plausible to me that gender differences for divorced and non-divorced individuals are also investigated.

In the migration context, I find it possible that the effect of a divorce/separation is different for male compared to female Heads - although I do not expect any gender differences in general (see discussion on Male in Part C, Section 4.1.1). Hence, although not explicitly surveyed in the literature, the interaction term MaleDivorce is included in the analysis without a predetermined expectation.

Interaction effects MaleAge, MaleEdu2, MaleEdu3

To the best of my knowledge, two-way interactions of gender and age, and gender and education, respectively, are only reported by Jianakoplos and Bernasek (1998). Unfortunately, their analysis relates only to the subgroup of unmarried individuals in the domain of financial risk-tolerance. For this subgroup Jianakoplos and Bernasek (1998) find significantly different effects of age and education for unmarried men and women.[263]

To control for this effect in the migration context, interaction terms MaleAge, MaleEdu2, and MaleEdu3 are included in the analysis without predetermined expectation.

4.1.2.2 Interactions with Single, Pair, Family, and Divorce

Besides the two-way interactions reported in the literature further interaction terms will be included. In this section two-way interactions based on dummies Single, Pair, Family, and Divorce each sequentially combined with Age, Edu2, and Edu3 are discussed. By including these interaction terms, I account for possibly different age- and education-effects on risk-attitudes depending on Head's family situation. Note that two-way interactions of Male with Single, Pair, Family, respectively have already been discussed in the previous section.

Further two-way interactions of age and education with Single, Pair, and Family

In the migration context, I expect that age- and education-effects differ for single-, pair- and family-movers for the following reason. In contrast to pair- and family movers, single-movers decide on their own about where to move, and do not need to consider other people's interest.

[263] Unfortunately, details on these findings are not reported by Jianakoplos and Bernasek (1998).

Therefore, I expect the effect of age and education to be more pronounced for single-movers, i.e., variables SingleAge, SingleEdu2, and SingleEdu3 are expected to have the same sign like the respective main effect on age or education, respectively. Vice versa, I expect age and education effects to be less pronounced for family- and pair-movers, i.e., variables PairAge, PairEdu2, PairEdu3, FamilyAge, FamilyEdu2, and FamilyEdu3 are expected to have the opposite sign than the respective main effects on age or education, respectively. Pair- and family-moves are considered separately because Heads of pair-moves might still exhibit different age- and education-effects than Heads of family-moves.

Further two-way interactions with Divorce: DivorceAge, DivorceEdu2, and DivorceEdu3

In line with the argumentation for the main effect of Divorce, divorced/separated Heads are expected to be significantly different from the rest of the sample therefore exhibiting a different age- and education-effect than other migrants.

I include DivorceAge, DivorceEdu2, and DivorceEdu3 without any predetermined expectation since the difference of divorced/separated Heads compared to all other migrants could result in both more and less pronounced age- and education-effects.

4.1.2.3 Interaction of age and education

Finally, I include the interaction effect of age and education (AgeEdu2, AgeEd3). While I expect a significant effect of these interactions, I have no clear expectation on the sign of the interaction. On the one hand, if I control for education, for example, by comparing the difference in risk-attitudes of a young and an old medical doctor to the difference of a young and an old craftsman, I expect the difference in risk-attitude due to age to be greater for the medical doctors. This suggests that the age-effect is more pronounced for higher levels of education, i.e., AgeEdu2 and AgeEdu3 have the same sign like age. On the other hand, if I control for age, for example, by comparing the difference in risk-attitudes of a young medical doctor and a young craftsmen to the difference of an old medical doctor and an old craftsman, I expect the difference in risk-attitude due to education to be more pronounced for the younger. This suggests that the education-effect is more pronounced for younger people, i.e., AgeEdu2 and AgeEdu3 have the same sign like education.

To sum up, I find it reasonable that the interaction effects AgeEdu2 and AgeEdu3 have the same sign like Age and the respective education variable (Edu2 or Edu3), but if Age and the respective education variable have opposing signs, I do not expect one effect to predominate the other.

4.2 Empirical methodology

To investigate the relation of risk-attitudes and the 32 independent variables discussed in the previous section, a suitable empirical methodology must be applied. This section explains why an extreme bounds analysis is the best methodological approach for my study (Part C, Section 4.2.1), how the extreme bounds analysis works in general (Part C, Section 4.2.2), and how it is applied to the data of my study (Part C, Section 4.2.3).

4.2.1 Reasons to apply extreme bounds analysis

Within the theoretical framework of my analysis, it is not clear which socio-economic characteristics relate to risk-attitudes. Therefore, several variables that possibly relate to risk-attitude either based on previous findings of the literature and/or based on economic considerations, were discussed in Part C, Section 4.1. But whether they truly relate to risk-attitudes in the migration context, and if so in which direction still needs to be proven. In other words, the data generating process is not known.

Usually, if this is the case, the "central method for selecting useful empirical models[s]"[264] is the general-to-specific modeling. It is used to decide which variables belong to the true model, and to figure out how these variables relate to the dependent variable.[265] Unfortunately, the approach of general-to-specific modeling is infeasible here for two reasons.

Problem 1 of general-to-specific modeling

For this study altogether 4,536 ways to estimate risk-attitudes have to be analyzed concerning their potential relation with 32 independent variables. Applying general-to-specific modeling results in different models for most ways to estimate risk-attitudes in the sense that one and the same independent variable will be included in some models and deleted in other. These different results for different ways to estimate risk-attitude make it impossible to draw general conclusions about the relation of independent variables on risk-attitudes irrespective of the way estimated. This is even more critical if one takes into account that the model resulting from general-to-specific modeling crucially depends on the algorithm used.[266]

---

[264] Campos, Ericsson, and Hendry (2005), p. 1.
[265] See Campos, Ericsson, and Hendry (2005), p. 1 and Krolzig and Hendry (2000), p. 1.
[266] See Pagan (1987) referred to in Campos, Ericsson, and Hendry (2005), p. 3.

Problem 2 of general-to-specific modeling

The time to run general-to-specific modeling for all 4,536 dependent variables would take more than 350 work days.

Solution to Problem 1

The solution to both problems is to run an extreme bounds analysis. In contrast to general-to-specific modeling, extreme bounds analysis does not delete variables from the model step by step, but concentrates on the distribution of regression coefficient of all independent variables.

Solution to Problem 2

The time to run an extreme bounds analysis with 32 independent variables and 4,536 dependent variables needs 93 work days to be finished.

4.2.2 Overview of extreme bounds methodology in general

The methodology of extreme bounds analysis applied in my study follows Sala-I-Martin (1997a, 1997b) who also includes a variant of the original approach going back to Leamer (1983). The objective of the extreme bounds analysis is to decide which independent variables have a robust influence on the dependent variable.

Speaking non-technically, extreme bounds analysis is a "global sensitivity analysis"[267].[268] More precisely, the idea behind it is to test the robustness of each independent variable by running a great number of regressions where the variable of interest is always kept in the regression. Only the other variables in the regression are altered in the way that (almost[269]) all subsets of independent variables that possibly belong to the model are included step by step. If the variable of interest shows similar regressions coefficients in all regressions, it is assumed to have a robust influence on the dependent variable.

[267] Leamer (1985), p. 308.
[268] What Leamer (1985) accurately called „global sensitivity analysis" was named "extreme bounds analysis" by one of his critics, McAleer, Pagan, and Volker (1985). The latter name became common in the aftermath.
[269] The idea is to test all combinations that make sense from an economic perspective and that can be estimated in a reasonable time period. The restriction from all combinations to only "almost" all is a result of defining fixed variables and determining the number of variables that are chosen among the non-fixed and not test variables to be entered in groups.

4.2.2.1 Defining regressions with fixed, non-fixed, and test variables

In order to run a sensitivity analysis on all independent variables using extreme bounds analysis, Sala-I-Martin (1997a, 1997b) recommends separating all independent variables into two groups, i.e., fixed and non-fixed independent variables[270]. Fixed variables are those independent variables that always appear in the regression because they were found to have a systematic impact on the dependent variable in most studies.[271] If they were missing in any regression, the model would clearly be misspecified.

The remaining independent variables are defined to belong to the second group of variables, i.e., non-fixed variables. From the non-fixed variables one after another is chosen to be the test variable for which robustness is tested.

To check robustness for each test variable, several regressions have to be run where each regression includes all fixed variables (vector $\boldsymbol{x_{fixed}}$) , the current test variable (x_{test}) from the non-fixed variables, and a subset of c non-fixed variable (vector $\boldsymbol{x_{non-fixed}}$) that are not the test-variable. Regressions are run until all possible combinations to choose c variables out of the remaining non-fixed variables have been estimated.

Each regression $m_{x_{test}} = 1, \dots, M_{x_{test}}$ on test variable x_{test} can be formalized by

$$\gamma = \beta_{0,m_{x_{test}}} + \boldsymbol{\beta}_{x_{fixed},m_{x_{test}}} \boldsymbol{x}_{fixed} + \beta_{x_{test},m_{x_{test}}} x_{test} + \boldsymbol{\beta}_{x_{non-fixed},m_{x_{test}}} \boldsymbol{x}_{non-fixed} \tag{16}$$

where γ denotes the dependent variable, $\beta_{0,m_{x_{test}}}$ the intercept of regression $m_{x_{test}} = 1, \dots, M_{x_{test}}$, $\boldsymbol{\beta}_{x_{fixed},m_{x_{test}}}$ a vector of coefficients of the fixed variables, $\boldsymbol{x}_{fixed}$ the vector of fixed variables, $\beta_{x_{test},m_{x_{test}}}$ the coefficient of the test variable, x_{test} the test variable, $\boldsymbol{\beta}_{x_{non-fixed},m_{x_{test}}}$ a vector of coefficients of a subset of c variables from the non-fixed variables that are not the test variable, $\boldsymbol{x}_{non-fixed}$ a subset of c variables from the non-fixed variables that are not the test variable.

Note that the number of independent variables in each regression must be kept constant for regressions run on one and the same test variable in order to be able to compare the regression results.

[270] Originally, Leamer (1983, 1985) and Leamer and Leonard (1983) refer to the non-fixed variables as “doubtful variables” while fixed variables are simply referred to as “variables which are certainly included in the equation” (Leamer and Leonard (1985), p. 307).
[271] See Sala-I-Martin (1997b), p. 8.

4.2.2.2 Testing the robustness of fixed and non-fixed variables

Although literature agrees on significance, sign, and robustness of the fixed variables, it is most often useful to check fixed variables again to gain confidence on previous findings for the following reasons: First, empirical results using one and the same model often differ when a different data set is used. This is due to samples that are often either not randomly sampled or do not include enough observations to be representative of the underlying population. Second, even if the same data set and model is used, a different estimation method could result in different findings. Third, for one and the same data set and estimation method, adding just one additional variable to the model could alter results. Fourth, a different specification of one and the same variable might also alter the results.

To test fixed variables for their robustness, one after another fixed variable is defined to be just like a regular test variable x_{test}, while the remaining fixed variables are kept in all regressions.[272] In other words, one after another fixed variable is redefined as test variable and the procedure as described in Part C, Section 4.2.2.1 is repeated.

Non-fixed variables are defined as such because previous findings in the literature are either inconclusive or ambiguous concerning its relation with the dependent variable in question. Consequently, these variables are tested for their robustness as described in the previous section.

4.2.2.3 Leamer's criterion of robustness

In order to judge the robustness of a certain independent variables on one dependent variable, two criteria exist. The first criterion applied by Sala-I-Martin (1997a, 1997b) is a variant of the original approach going back to Leamer (1983, 1985) and Leamer and Leonard (1983) that was first applied by Levine and Renelt (1992). Following the convention in the literature, I refer to the criterion discussed in this section as Leamer's criterion of robustness.

The idea of Leamer is to derive a lower and upper extreme bound from all regressions run on a test variable. If upper and lower extreme bounds have the same (different) sign, the corresponding test variable is defined as robust (fragile). The implementation of Leamer's criterion works as follows:

[272] See Sala-I-Martin (1997b), p. 13.

1st step: Running all regressions, one finds for each test variable x_{test} and each regression $m_{x_{test}} = 1, .., M_{x_{test}}$ a regression coefficient $\beta_{x_{test}, m_{x_{test}}}$ and a corresponding standard error $\sigma_{x_{test}, m_{x_{test}}}$ from which the corresponding 95%[273] confidence interval can be derived as follows

$$\left[\beta_{x_{test}, m_{x_{test}}} - 2\sigma_{x_{test}, m_{x_{test}}}; \beta_{x_{test}, m_{x_{test}}} + 2\sigma_{x_{test}, m_{x_{test}}}\right]. \tag{17}$$

where $\beta_{x_{test}, m_{x_{test}}}$ denotes the regression coefficient of test variable x_{test} in regression $m_{x_{test}}$, $\sigma_{x_{test}, m_{x_{test}}}$ the robust standard error of the regression coefficient relating to test variablex_{test} in regression $m_{x_{test}}$.

Note that, although not explicitly mentioned in any paper of the literature, this definition implicitly assumes a normal distribution of regression coefficients.[274]

2nd step: From these confidence intervals the lower extreme bound is defined as minimum of $\beta_{x_{test}, m_{x_{test}}} - 2\sigma_{x_{test}, m_{x_{test}}}$ found among all $M_{x_{test}}$ regressions run on test variable x_{test}. Vice versa, the upper extreme bound is defined as maximum of $\beta_{x_{test}, m_{x_{test}}} + 2\sigma_{x_{test}, m_{x_{test}}}$ found among all $M_{x_{test}}$ regressions run on test variable x_{test}.[275]

3rd step: If lower and upper extreme bounds have the same (different) sign, the corresponding dependent variable is defined as robust (fragile).

Leamer's criterion is very strict since it amounts to saying that if (i) only one single coefficients of all $M_{x_{test}}$ coefficients becomes either insignificant or (ii) has another sign than all other coefficients, test variable x_{test} is classified as fragile.[276] The reasons for this are as follows: Concerning the first case (only one coefficient is insignificant), note that confidence intervals of insignificant coefficient always have a negative lower boundary and a positive upper boundary. This means, in the case when regression coefficients of all other regressions have the same sign and are significant (i.e., their upper and lower boundary of the confidence intervals are strictly positive or negative, respectively), at least

[273] Although reported as 95% confidence interval in Levine and Renelt (1992), p. 944, the confidence interval precisely amount to 95.4%.
[274] See for example, Levine and Renelt (1992), p. 944. The original idea of Leamer and Leonard (1983) was that if the difference between the minimum and maximum estimate of the regression coefficient over all regressions is small in relation to the sampling uncertainty, all regressions result in the same inference. Although Leamer and Leonard (1983) give several examples about how this relation can be defined, they leave the concrete definition open to the researcher. Therefore, the interval of Equation (17) can also be considered as any criterion that tries to relate minimum and maximum estimates to sampling uncertainty if it is refrained from the definition as 95% confidence interval.
[275] Note that the lower and upper extreme bounds need not necessarily stem from one and the same regression.
[276] See Sala-I-Martin (1997a), p. 178.

one boundary of the insignificant confidence interval becomes an extreme bound. Therefore, lower and upper extreme bounds have different signs and the test variable is classified as fragile. Concerning the second case (only one coefficient has a sign different from all other coefficients), two scenarios must be distinguished. If the coefficient is insignificant, the same argument as for case one holds. If it is significant, both boundaries of the confidence interval will have the same sign, while for all other coefficients at least one boundary of the confidence interval has a different sign. Ergo, lower and upper bounds of the extreme bounds analysis have different signs and the variable is classified as fragile.

To sum up, robustness in Leamer's sense always means the test variable is significantly related to the dependent variable where the sign of the coefficient is the same as the sign of both, upper and lower extreme bounds. The opposite does not apply. Fragility does not necessarily mean insignificant. Therefore, if a test variable is defined as fragile by Leamer's criterion, no statement on significance and sign can be derived.

4.2.2.4 Sala-I-Martin's criteria of robustness

Since Leamer's criterion is very strict, it is not surprising that many studies applying Leamer's robustness test did not find much support for the variables tested. Instead of assigning the label "fragile" to them, Sala-I-Martin (1997a, 1997b) offers another possible explanation: Maybe the test is "too strong for any variable to really pass it."[277] He argues that if the distribution of the true coefficient of a variable has "some positive and some negative support"[278], it is not surprising that a coefficient with another sign is found if enough regressions are run.[279]

Consequently, the alternative criterion of robustness Sala-I-Martin (1997a, 1997b) offers, is to assign some confidence level instead of using a binary label (robust/fragile). He derives a distribution of the regression coefficient where a test variable is defined as robust if the probability for a certain sign (positive or negative) under this distribution amounts to at least 95%. The implementation of Sala-I-Martin's criterion works as follows:

[277] Sala-I-Martin (1997b), p. 4.
[278] Sala-I-Martin (1997b), p. 4.
[279] See Sala-I-Martin (1997b), p. 4.

1st step: He assumes normally distributed coefficients across the $M_{x_{test}}$ regressions of test variable x_{test}.[280] Parameter values mean $\overline{\beta_{x_{test}}}$ and variance $\overline{\sigma_{x_{test}}}^2$ of the distribution are both estimated as weighted average of all $M_{x_{test}}$ point estimates $\beta_{x_{test},m_{x_{test}}}$ and $\left(\sigma_{x_{test},m_{x_{test}}}\right)^2$, respectively

$$\overline{\beta_{x_{test}}} = \sum_{m_{x_{test}}=1}^{M_{x_{test}}} w_{m_{x_{test}}} \beta_{x_{test},m_{x_{test}}} \quad \text{and} \quad \overline{\sigma_{x_{test}}}^2 = \sum_{m_{x_{test}}=1}^{M_{x_{test}}} w_{m_{x_{test}}} \sigma_{x_{test},m_{x_{test}}}{}^2 \qquad (18)$$

where $\overline{\beta_{x_{test}}}$ denotes the average regression coefficient of test variable x_{test}, $w_{m_{x_{test}}}$ the weight of regression $m_{x_{test}}$ testing test variable x_{test}, $\overline{\sigma_{x_{test}}}^2$ the average variance of test variable x_{test}.

The weighting is performed to account for the fact that some regression models are closer to the true model than others. Therefore, regression estimates are weighted by their likelihoods in order to give higher weights to regression model with higher explanatory power. Weights are constructed by

$$w_{m_{x_{test}}} = \frac{\mathrm{L}_{m_{x_{test}}}}{\sum_{m_{x_{test}}}^{M_{x_{test}}} \mathrm{L}_{m_{x_{test}}}} \qquad (19)$$

where $\mathrm{L}_{m_{x_{test}}}$ denotes the likelihood of regression $m_{x_{test}}$ testing test variable x_{test}.

Note that likelihoods are not biased due to varying numbers of independent variables because the number of independent variables is held constant over all regressions on a certain test variable. Still, likelihoods might be biased due to endogenous independent variables that might results in a spurious better fit of the model.[281] In this case, all or almost all weight might be put on only a few regressions which systematically bias the result. To control for this problem, average mean and variance are also estimated as arithmetic average $\left(w_{m_{x_{test}}} = \frac{1}{M_{x_{test}}}\right)$. – In the further course of my analysis, I refer to this approach as "unweighted" in contrast to "weighted".

2nd step: Once mean and variance of the regression coefficient are estimated, the corresponding normal distribution can be used to estimate the probability that the coefficient has a certain sign. Technically, the probability for a certain sign of the coefficient can be derived by dividing the area under the density function in two regions at point zero. The area below zero relates to the

[280] Sala-I-Martin (1997a, 1997b) also operates under a non-normal assumption, but this approach is not implemented here because coefficients estimated by Maximum-Likelihood algorithm are assumed to follow a normal distribution.
[281] Endogenous independent variables correlate with the error term so that the explanatory power of the regression is overestimated.

probability that the test variable has a negative influence on the dependent variable. Conversely, the area above zero relates to the probability of a positive coefficient. Sala-I-Martin (1997a, 1997b) calls the greater of the two areas – no matter if it is below or above zero – $CDF(0)$ which naturally lies between 0.5 and 1. Based on this figure, Sala-I-Martin (1997a, 1997b) defines a variable to be robust if 95% of the cumulative density function lies on one side of zero, i.e., $CDF(0) \geq 0.95$.

To sum up, robustness in Sala-I-Martin's sense means that the probability for a certain sign of the regression coefficient amounts to at least 95%, i.e., a statement on sign and robustness of the test variable can be made, but a statement on significance cannot be derived.

4.2.3 Application of extreme bounds analysis to my variables

4.2.3.1 Problems and solutions in applying extreme bounds methodology to my study

Problem 1: Several dependent variables

In my study the 4,536 ways to estimate risk-attitude have to be analyzed concerning their relation with various socio-economic characteristics. This means, in contrast to the general methodology of the extreme bounds analysis described in Part C, Section 4.2.2, which relates to only one single dependent variable, I have 4,536 dependent variables in my study.

Solution to Problem 1: Extreme bounds analysis run separately for each dependent variable

I decide to run an extreme bounds analysis separately for each dependent variable for two reasons: First, is not clear whether the 4,536 dependent variables relate to the same decision problem. For example, taking the decision to migrate to another state just for a single year (planning period of one year) might be another decision problem than deciding about where to migrate for the rest of life (planning period of time until reaching life expectancy). If different dependent variables do not relate to the same decision problem, they are not comparable in economic terms. Second, even if the dependent variables were comparable in economic terms, the estimation problem remains. In other words, different dependent variables relate to different ways of both, the way data input is gained and the way risk-attitudes are estimated from this data input, and it is not clear which approach is the right one.

Problem 2: Binary dependent variables

The approach of Sala-I-Martin (1997a, 1997b) as presented in Part C, Section 4.2.2 assumes that the dependent variable is scaled metrically. This is not suitable for my study because my dependent variables are binary coded.

Solution to Problem 2: Binary logistic regression

The solution is to apply a binary logistic regression instead of an ordinary linear regression. This can be done without any problems because the dependent variable of a logistic regression, the log of the odds of being risk-averse (also called logit), can be expressed as some linear combination of the independent variables as follows

$$\ln\left(\frac{\lambda}{1-\lambda}\right) = \beta_{0,m_{x_{test}}} + \boldsymbol{\beta}_{\boldsymbol{x_{fixed}},m_{x_{test}}}\,\boldsymbol{x_{fixed}} + \beta_{x_{test},m_{x_{test}}}\,x_{test} + \boldsymbol{\beta}_{\boldsymbol{x_{non-fixed}},m_{x_{test}}}\,\boldsymbol{x_{non-fixed}} \tag{20}$$

where $\ln\left(\frac{\lambda}{1-\lambda}\right)$ denotes the logit of the probability of being risk-averse λ.

Problem 3: Independent variables of my study cannot be freely combined

In the general methodology discussed in Part C, Section 4.2.2, it is implicitly assumed that all independent variables can enter the regression in any possible combination. This assumption is not critical if two prerequisites are met. First, all categorical variables with k levels must be transformed to no more than k-1 dummies so that the dummy trap is avoided in any case. This is a problem if the effect of all k categories is analyzed separately and no category should be excluded from the analysis. The second prerequisite is that all independent variables must be either exclusively main effects or exclusively interaction effects. The reason is that the interpretation of interaction effects depends on whether their main effects are also entered in the regression or not. This means, if independent variables consist of main effects and interaction effects, and these variables are freely combined, then interaction effects will sometimes be included in the regression without its main effects, and sometimes together with its main effects. Consequently, the regression coefficients of the interaction term are not comparable anymore since they do not measure the same thing in all regressions, and Leamer's and Sala-I-Martin's criteria of robustness cannot be applied anymore.

Both presumptions are violated in my study. First, I have variables such as the dummies relating to singe-, pair-, and family-moves all of which I wish to include in the analysis separately. Second, I wish to analyze a great variety of variables including both interaction effects and corresponding main effects. Therefore, the independent variables of my study cannot be freely combined. This raises the question on how to combine my 32 independent variables of my study when running the extreme bounds analysis.

Solution to Problem 3: General rules to combine independent variables

To solve Problem 3, two rules on how to combine independent variables are applied: The first rule simply states that the dummy trap must be avoided. This is important, for example, if Single is the test variable and a subset of $c = 2$ variables has to be entered. In this case, Pair and Family must not be included at the same time.

The second rule on how to combine independent variables relates to interaction effects that are only allowed to be entered together with their main effects. This is done for two reasons: First, interaction terms allow for different slopes of regression lines of different categories.[282] If the main effects of the interaction term are not contained in the regression, only the slopes of the regression lines would be allowed to differ, but the intercept would be the same for all regression lines.[283] Second, if one main effect was missing, the interaction coefficient might be biased because it additionally includes the missing main effect - which would have been captured by the coefficient of the main effect otherwise.

Problem 4: No standard software existent

Although several software applications on extreme bounds analysis exist, the two rules discussed in the solution to Problem 3 make it impossible to use standard software. To the best of my knowledge there is also no user written program available that accounts for my rules.

Solution to Problem 4: Programming of own algorithm

In order to be to able run an extreme bounds analysis that accounts for my rules on how to combine independent variables, I program a source code that runs the respective regressions in STATA 11 and exports the results to comma-separated values data files (csv). These are imported to Excel where Leamer's and Sala-I-Martin's criteria of robustness are finally applied as described in Part C, Sections 4.2.2.3 and 4.2.2.4.

---

[282] See Stocker (2013), p.6.
[283] See Stocker (2013), p.7 f.

4.2.3.2 Choosing fixed, non-fixed, and test variables

Variables on gender, age, and education are defined as fixed variables in my study for the following reasons: Although a study on risk-attitude concerning the migration decision does not exist so far, a great body of literature is concerned with determinants of risk-attitudes in other non-migration domains. Almost all studies of this kind find gender, age, and education to have a robust significant impact on risk-attitude.[284] In order to avoid a misspecification due to omitting important variables, variables on gender, age, and education are defined as fixed variables in my study.

Note that the education variable consists of three levels. In order to avoid the dummy trap, two dummies on education are entered as fixed variables in all regressions where the lowest education level is the reference category. Consequently, the number of fixed variables entered in this study is four, i.e., Male, Age, Edu2, and Edu3 as defined in Part C, Section 4.1. The remaining 28 independent variables of Part C, Section 4.1 are defined as non-fixed variables.

While testing robustness of non-fixed variables is obligatory, testing fixed variables is not necessary for any study. Still, the previous empirical literature (i) relates to risk-attitudes in other non-migration domains, and (ii) has shown that risk-attitudes are domain-specific. Therefore, previous findings cannot necessarily be transferred to the domain of migration. Consequently, I decide to additionally test the fixed variables of my study to gain confidence on their significance, sign, and robustness in relation with risk-attitude in the migration context.

The testing of fixed and non-fixed variables is performed as described in Part C, Section 4.2.2.2.

4.2.3.3 Choosing the number of variables to be entered in sets (c)

Although no rule exists on how to choose the number of variables to be entered in sets (c), it is a crucial parameter in the extreme bounds analysis. Basically, there are two conflicting arguments. On the one hand, the model to be estimated should reach a certain minimum level of explanatory power. This means, c should not be too low. On the other hand, a higher value of c soon results in very high numbers of regressions to be run, which soon takes up years to be estimated with the current computer technology. The latter argument is a real issue in this study. Recall, that 32 independent variables have to be tested for each of the 4,536 dependent variables on risk-attitude.

Usually, the trade-off can easily be solved by choosing the greatest value of c that is still manageable given the current computer capacity. This requires that the number of regression to be run is known

[284] For a detailed discussion of findings in the literature see Part C, Section 4.1.

in order to be able to estimate the computer capacity needed. Generally, the number of regressions to be run for each test variable in the extreme bounds analysis $\left(M_{x_{test}}\right)$ can easily be calculated by the number of the remaining non-fixed variables that are not the test variable (rnf), and the number of variables to be entered in sets (c) as follows

$$\binom{rnf}{c} = \frac{rnf!}{c! \cdot (rnf - c)!} \tag{21}$$

where rnf denotes the number of remaining non-fixed variables that are not the test variable, and c the number of variables to be entered in sets.

For example, if four of the 32 independent variables are fixed variables, one variable is the test variable, and the remaining 27 non-fixed variables are entered in sets of two $(c = 2)$, this results in $\binom{27}{2} = \frac{27!}{2! \cdot (27-2)!} = 351$ regressions to be run for each non-fixed test variable and each dependent variable.

Unfortunately, this calculation is wrong in my study. Equation (21) only indicates the theoretical maximum number of regressions to be run if all variables can be freely combined with each other, but this is not the case in my study (see Part C, Section 4.2.3.1, Problem 3 and the related solution). Therefore, in my study the number of regressions $M_{x_{test}}$ to be run for test variable x_{test} cannot be estimated using Equation (21), but rather has to be carefully calculated separately for each test variable taking the two rules to combine non-fixed variables into account (for the rules see Part C, Section 4.2.3.1, Problem 3 and related solution).

Taking the two rules of how to combine non-fixed variables into account, a reasonable compromise for my study is to allow non-fixed variables to enter regressions in sets of two $(c = 2)$. For most test variables this results in seven explanatory variables, i.e., four fixed-variables, the test variable and a subset of two variables out of the remaining non-fixed variables. A value of $c = 2$ is chosen for two reasons. First, regressions with seven explanatory variables can be expected to reach a satisfying level of explanatory power. A reasoning shared by Sala-I-Martin (1997a, 1997b), who also uses seven independent variables for his analysis. Second, the time to compute all regressions accounts for acceptable 93 work days. If a value of $c = 3$ was chosen, it would take more than a work year to run all regressions. The total number of regressions to be run and how variables are combined in detail are discussed in the next section.

4.2.3.4 Over 12 million regressions to be run

If non-fixed variables that are not the test variables are entered in sets of two $(c = 2)$, this result in a total of more than 12 million regression to be run for my study. Table 35, p. 191 gives an overview of the number of regressions to be run separately for each independent variable and one dependent variable $\left(M_{x_{test}}\right)$ if the two rules on how to combine non-fixed variables (i.e., avoiding the dummy trap, and enter interaction effects only together with their main effects, see Part C, Section 4.2.3.1, Problem 3 and related solution) are taken into account. Where these numbers come from is discussed in detail below.

	Test variable x_{test}	**Number of regressions $\left(M_{x_{test}}\right)$ run for test variable x_{test}**
Category 1: Fixed variables		
1	Male	82
2	Age	82
3	Edu2	82
4	Edu3	82
Category 2: Non-fixed main effects not included in any interaction		
5	AgeSquared	71
6	Number of family members in wave before the move	71
7	Number of family members in wave after the move	71
Category 3: Non-fixed main effects included in some interactions		
8	Single	116
9	Pair	116
10	Family	116
11	Divorce	117
Category 4: Non-fixed interaction effects where both main effects are fixed variables		
12	AgeMale	71
13	AgeEdu2	71
14	AgeEdu3	71
15	MaleEdu2	71
16	MaleEdu3	71
Category 5: Non-fixed interaction effects where only one main effect is a fixed variable		
17	SingleMale	102
18	SingleAge	102
19	SingleEdu2	102
20	SingleEdu3	102

21	PairMale	102
22	PairAge	102
23	PairEdu2	102
24	PairEdu3	102
25	FamilyMale	102
26	FamilyAge	102
27	FamilyEdu2	102
28	FamilyEdu3	102
29	DivorceMale	103
30	DivorceAge	103
31	DivorceEdu2	103
32	DivorceEdu3	103
Sum of regressions for each dependent variable		**2,997**[285]

Table 35: Overview of 32 independent variables and corresponding number of regression $M_{x_{test}}$ to be run for one dependent variable if variables are entered in sets of two $(c = 2)$.
Source: Own calculations.

For each of the five categories, the amount of explanatory variables in each regression and the number of regressions to be run can be calculated as follows:

Category 1: Test variables that are fixed variables

The regressions to test fixed test variables always consists of the four fixed variables (one after another interpreted as test variable) plus a subset of two variables out of the 28 non-fixed variables. Ergo, each regression has six explanatory variables. The 82 regression to be run for fixed test variable 1 from Table 35, p. 191 result from varying combinations as follows: First, variables 5 to 16 can be freely combined in pairs of two. All variables are either main effects themselves or both of their main effects are fixed variables that are included in the regression anyway (second rule). Furthermore, a dummy trap (first rule) cannot occur since only two variables are entered at the same time and neither Single, Pair, nor Family is a fixed variable. Therefore, variables 5 to 16 can be combined in 12!/[2!·(12-2)!]=66 ways. Second, according to the second rule, variables 17 to 32 must be combined with a specific other variable to have all their main effects included in the regression: Variables 17 to 20 can only be combined with Single, variables 21 to 24 can only be combined with Pair, variables 25 to 28 can only be combined with Family, and variables 29 to 32 can only be combined with Divorce. This results in further 16 combinations. Therefore, a total of 82 regressions must be run for test variable 1 from Table 35, p. 191.

---

[285] Since regressions to be run to test the fixed variables on gender, age, and education are identical. This means, if all regressions on Male are run, the same regressions can be used for testing Age, Edu2, and Edu3, the number of regressions to be run is reduces to 2,751.

The same logic holds for regressions to test fixed variables 2 to 4 from Table 35, p. 191. Note that all regressions always include the four fixed variables no matter which of the fixed variables is considered to be the test variable, and 82 different combinations of the non-fixed variables are identical for all fixed variables. Therefore, the 82 regressions are only run once, but their results are evaluated four times - once for each of the four fixed test variables.

Categories 2 and 4: Test variables that are non-fixed variables that are either main effects not included in any interaction or interaction effects where both main effects are fixed variables

Regressions to test this group of variables always consist of four fixed variables, the test variable and a subset of two variables out of the remaining 27 non-fixed variables. Ergo, each regression has seven explanatory variables. Exemplary for test variable AgeSquared, the 71 regressions result from varying combinations as follows: First, note that Age as main effect of AgeSquared is a fixed variable already included in all regressions. Second, variables 6 to 16 can be freely combined in pairs of two for reasons already discussed in the first example (combinations for test variables of Category 1). This results in 11!/[2!·(11-2)!]=55 combinations. Third, like in the first example, variables 17 to 32 all have a predetermined second variable which results in further 16 combinations. In sum, 71 regressions must be run to test the robustness of test variable AgeSquared.

For all other test variables of Categories 2 and 4 from Table 35, p. 191, non-fixed variables are combined using the same logic. Hence, for them the same number (71) of regressions must be run.

Category 3: Test variables that are non-fixed main effects included in some interaction effects

All regressions to be run to test this group of variables (Single, Pair, Family, Divorce) include seven explanatory variables: four fixed variables, a test variable and a subset of two non-fixed variables. Exemplary for this group of test variables, consider test variable Single and its 116 regressions to be run. First, of all possible combinations among variables 5 to 7 and 9 to 16 (i.e., 11 variables) only the combination of variables Pair and Family is not allowed in order to avoid the dummy trap. Therefore, variables 5 to 7 and 9 to 16 can be combined in 11!/[2!·(11-2)!] – 1 = 54 ways. Second, since all main effects of interaction terms SingleMale, SingleAge, SingleEdu2, and SingleEdu2 are included in the regression either as test variable (Single) or as fixed variable (Male, Age, Edu2, Edu3), each of the interaction effects can be combined with variables 5 to 7, and 9 to 16. This results in further 4 · 11 = 44 combinations. Third, interaction terms SingleMale, SingleAge, SingleEdu2, and SingleEdu2 can also be combined with each other, resulting in additional 4!/[2!·(4-2)!] = 6 combinations. Fourth, variables

21 to 24 can only be combined with Pair, variables 25 to 28 can only be combined with Family, and variables 29 to 32 can only be combined with Divorce. This results in further 12 combinations.

For all other test variables of Category 3 from Table 35, p. 191 non-fixed variables are combined using the same logic, resulting in the same number of regression to be run with one exception: For test variable 11 the dummy trap as discussed under the first point in this paragraph does not exist. Therefore, the combination of variables Pair and Family is additionally possible resulting in 117 regressions to be run for test variable 11 from Table 35, p. 191.

Category 5: Test variable that are non-fixed interaction effects where only one main effect is a fixed variable

To test this group of variables, the test variable's main effect that is no fixed variable must be additionally included in all regressions (second rule). This results in eight explanatory variables of each regression: four fixed variables, the test variable itself, the non-fixed main effect of the test variable, and a subset of two variables out of the remaining 26 non-fixed variables. Exemplary for this group of test variables, consider test variable DivorceMale and its 103 combinations as follows: First, have in mind that Divorce as the non-fixed main effect of the test variable is additionally included in all regressions. Second, variables 5 to 16 (except for variable 10, Divorce) can freely be combined because they are either main effects themselves (variables 5 to 11) or interactions where both main effects are fixed variables already included in the regression (variables 12 to 16). Therefore, neither the first nor the second rule put restrictions on combining these variables which results in $11!/[2!\cdot(11-2)!]=55$ combinations. Third, variables 17 to 28 are entered together with their predetermined second variables (12 combinations) as describes in the first example (Category 1 from Table 35, p. 191). Fourth, variables 30 to 32 (DivorceAge, DivorceEdu2, DivorceEdu3) can be freely combined with variables 5 to 16 (again except variable 10 Divorce) because their non-fixed main effect Divorce has already been included in the regression, and variables 5 to 16 are not restricted by neither the first nor the second rule. This results in further $3 \cdot 11 = 33$ combinations. Fifth, variables 30 to 32 can be combined among each other (3 combinations).

The same logic holds for interaction effects DivorceAge, DivorceEdu2, and DivorceEdu3. For variables 17 to 28 one regression less is estimated due to the dummy trap. For example, if SingleMale is the test variable, Single is always included in the regression as non-fixed main effect. Therefore, the combination of variables Pair, Family is invalid in order to avoid the dummy trap. This results in one regression less compared to interactions of Divorce and other non-fixed main effect.

<u>Total number of regressions to be run for all 4,536 dependent variables on risk-attitude if variables are entered in pairs of two $(c = 2)$</u>

To test the robustness of all 32 independent variables concerning one dependent variable, 2,997 regressions must be run. If we account for the identical regressions for each fixed test variables, the number reduces to 2,751 regressions. Consequently, for the 4,536 dependent variables on risk-attitude that must be analyzed in this study a total of 12,478,536 regressions must be run. The time needed to gain regression coefficients for the study at hand took 93 work days to be finished.

4.3 Empirical results on socio-economic characteristics

4.3.1 Methodology of interpretation

The objective of the empirical analysis is to reach a clear-cut statement on significance, sign, and robustness for each of the 32 independent variables on socio-economic characteristics concerning their relation to the depended variable on risk-attitudes, irrespective of the way risk-attitudes are estimated.

Unfortunately, the results of the 4,536 extreme bounds analyses run separately for each dependent variable show that one and the same independent variable is related significantly positive to one dependent variable, significantly negative to another, and not significantly related to a third. This is not surprising if it is recalled that different dependent variables might relate to (i) different decision problems, (ii) different ways to gain data input, and (iii) different estimation procedures (see Problem 1 and related solution in Part C, Section 4.2.3.1). Certainly, the desired clear-cut statement irrespective of the way the dependent variables is estimated does not exist. Therefore - and because printing all results of the 4,536 ways to estimate risk-attitudes would fill more than 8,500 pages full of tables - I refrain from reporting detailed results here, but kindly ask the interested reader to refer to folder „\Alpha\(10) Ergebnisse sammeln" on the data storage that comes along with this study. Instead, the following methodology is applied to interpret the empirical results.

<u>Central methodology of interpreting empirical results</u>

The different results for different dependent variables raise the question of what can be learned from such an analysis of several dependent variables. To answer this question, I carry out for each independent variable a joint interpretation over all 4,536 dependent variables. The central methodology I apply to achieve the joint interpretation is as follows: For each independent variable, I calculate the percentage of the dependent variables for which a certain criterion (e.g., Leamer's or Sala-I-Martin's weighted/unweighted criterion) is met.

Details on the central methodology

In detail, the following methodology is applied for each independent variable:

- **Sample size and (quasi-) complete separation**: First, I check whether the sample size is high enough to derive a statistical statement on that independent variable. Sample size in this context relates to the number of dependent variables for which an extreme bounds analysis on the independent variable in question could have been run. Ideally, results on all 4,536 dependent variables are available. If missing values occur (i.e., the sample size is smaller than 4,536), this is due to complete or quasi-complete separation. (Quasi-) Complete separation means that the independent variable (almost) perfectly predicts the outcome of the dependent variable. (Quasi-) Complete separation most certainly is a result of the sample size being too small. In such a situation of (quasi-) complete separation, a maximum likelihood estimator of the corresponding independent variable does not exist. Hence, there are no regression coefficients for which Leamer's or Sala-I-Martin's weighted/unweighted criterion of robustness can be applied. Therefore, dependent variables that exhibit (quasi-) complete separation are excluded from the statistical analysis. Vice versa, all dependent variables that do not exhibit (quasi-) complete separation are included in the statistical analysis of my study.

  An example may help to demonstrate the situation of complete and quasi-complete separation: Table 36, p. 195 gives the cross tabulation of an exemplary dependent variable on risk-attitude and the independent variable Edu2 (i.e., dummy for people with highest education level high school diploma or associate's degree). In this example, quasi-complete separation occurs. That is, all people with education levels other than high school diploma/associate's degree (Edu2=0) are risk-averse. This means the probability of being risk-averse is 100% if the person has a highest education level other than high school diploma/associate's degree (Edu2=0). If the person has a high school diploma or an associate's degree (Edu2=1), the probability of being risk-averse is still 90% (i.e., 83 divided by 92). In contrast to quasi-complete separation, complete separation would mean that all people with education level Edu2 (Edu2=1) are risk-seeking and all people with another education level (Edu2=0) are risk-averse. Put differently, the value of the independent variable Edu2 perfectly predicts the outcome of the dependent variable.

	Risk-aversion	**Risk-seeking**	**Total**
Highest education level high school diploma/ associate's degree (Edu2=1)	83	9	92
All other levels of education (Edu2=0)	57	0	57
Total	140	9	149

Table 36: Example of quasi-complete separation for Edu2 (dummy on highest education level of high school/associate's degree) and dependent variable 1 (Fam_WeiPooLP2LifL1Ad2Ed1).

For all dependent variables that do not exhibit complete or quasi-complete separation, a statistical statement on significance, sign, and robustness is derived step by step as follows:

- **Robustness and significance based on Leamer's criterion:** Concerning robustness, Leamer's criterion is interpreted as follows: The higher the percentage of dependent variables for which an independent variable is found to have a robust influence by Leamer's criterion, the higher the probability that the independent variable is truly related to the probability of being risk-averse in the migration context. Recall that robustness according to Leamer's criterion is equivalent to saying that the influence is significant.

  To classify an independent variable as *robust by Leamer's criterion* irrespective of the way risk-attitudes are estimated (i.e., over all 4,536 dependent variables), I require the independent variable to be robustly related by Leamer's criterion to at least 75% of the dependent variables statistically investigated.[286]

- **Robustness based on Sala-I-Martin's criterion:** Concerning robustness, Sala-I-Martin's weighted/unweighted criterion is interpreted as follows: The higher the percentage of dependent variables for which an independent variable is found to have a robust influence by Sala-I-Martin's weighted/unweighted criterion, the higher the probability that the independent variable is truly related to the probability of being risk-averse in the migration context.

  To classify an independent variable as *robust by Sala-I-Martin's criterion* irrespective of the way risk-attitudes are estimated (i.e., over all of the 4,536 dependent variables that are empirically analyzed), I require the independent variable to be robustly related by both Sala-I-Martin's weighted and unweighted criteria to at least 75% of the dependent variables statistically investigated.[287]

- **Tendency for a sign based on Leamer's and Sala-I-Martin's criteria**: Both Leamer's and Sala-I-Martin's criteria of robustness also indicate a sign of the relation between an independent

[286] Caveat: Leamer's criterion of robustness as described in Part C, Section 4.2.2.3 refers to only one single dependent variable. It requires that the independent variable shows a significant relation of the same direction in 100% of the regressions. In contrast to this approach, I am now interested in interpreting the relation of an independent variables for all dependent variables not only one dependent variable. Therefore, I require at least 75% of the dependent variables to meet Leamer's 100%-criterion.

[287] Caveat: Sala-I-Martin's weighted and unweighted criterion of robustness as described in Part C, Section 4.2.2.4 refers to only one single dependent variable. It requires that the probability for a certain sign of the coefficient of the independent variable is at least 95%. In contrast to this approach, I am now interested in interpreting the relation of an independent variables for all dependent variables not only one variables. Therefore, I require at least 75% of the dependent variables to meet both Sala-I-Martin's weighted and unweighted 95%-criterion.

variable and the probability of being risk-averse as follows: The higher *the percentage of robust relations* that are positive (negative), the higher the probability that the independent variable is positively (negatively) related to the probability of being risk-averse. This approach can be applied to robust relations by Leamer's and both Sala-I-Martin's weighted and unweighted criteria.

Irrespective of robustness, the sign of the relation can also be judged by Sala-I-Martin's weighted/unweighted criterion as follows: The higher *the percentage of all relations* (irrespective of their robustness) where the probability of a positive (negative) sign is greater than 50%, the higher the probability that the independent variable is positively (negatively) related to the probability of being risk-averse. This approach is chosen because Sala-I-Martin's weighted/unweighted criterion requires an arbitrarily chosen minimum probability of 95% to consider a relation as robust. Relations that, for example, only exhibit a probability of 94% for a certain sign would be neglected.

In order to judge the sign of the relation, there are five percentages (percentages referring to robust relations applying Leamer's criterion, and percentages referring to robust/all relations applying Sala-I-Martin's weighted/unweighted criterion) that have to be jointly interpreted in order to derive a statistical statement on the sign of the relation between the independent variable and the probability of being risk-averse in the migration context. Note that the number of percentages to be interpreted may be reduced to less than five when first, Leamer's or Sala-I-Martin's weighted/unweighted criteria do not find any robust relations and the corresponding percentage referring to robust relations does not exist. Second, the number of five percentages might be reduced when percentages amount to exactly 50%. Since 50% neither indicates a positive nor a negative sign, percentages that amount to 50% are not considered for interpretation. Consequently, those percentages that are considered for interpretation are jointly interpreted as follows:

Since the percentages are at best tendencies for the true relation of the independent variable on the probability of being risk-averse, I demand all percentages available to point in the same direction in order to regard them as indicating **"a vague tendency for a positive/negative sign"**. If all percentages available point in the same direction and at least one of the percentages available additionally reaches at least a 75% majority, I regard the results as indicating **"a strong tendency for a positive/negative sign"**. In contrast, if only one percentage points in another direction than all other available percentages, I regard the results as indicating **"no tendency for a sign"**.

Sensitivity analysis for subgroups of dependent variables

The 4,536 dependent variables in my study result from a combination of competing solutions of problems that occurred when the theoretical migration decision model was applied to real data. For example, the problem of how to measure risk was solved by the competing solutions to the risk-measure problem variance and semi-variance (see Part C, Table 31, p. 160 for an overview of all competing solutions). It is therefore possible that empirical results systematically differ for subgroups of dependent variables that relate to competing solutions. In this case the joint interpretation over all 4,536 dependent variables would be is misleading. To investigate whether empirical results systematically differ for competing solutions of one and the same problem, a sensitivity analysis is run.

An example illustrates the idea of the sensitivity analysis: To check the sensitivity of the results with respect to competing risk-measures, the 4,536 dependent variables are divided into (i) a subgroup of dependent variables that result from risk-measure variance, and (ii) a second subgroup of dependent variables that result from risk-measure semi-variance. The central methodology is then applied separately to both subgroups and results are compared. In the case statements on significance, sign, and robustness derived for each subgroup are different, results are found to be sensitive to different risk-measures. Note that the subgroups of dependent variables are identical except for the risk-measure applied. Their comparison therefore works like a ceteris paribus analysis in that only the competing solution is modified and all other independent variables are not modified.

Since I want to test the sensitivity not only with respect to competing risk-measure, but with respect to all competing ways to solve one and the same problem that occurred when the theoretical model was applied to real data (see Table 31, p. 160 for an overview), I perform a ceteris paribus sensitivity analysis of subgroups of dependent variables as follows: I compare first, subgroups relating to competing decision problems (i.e., family versus personal income, different planning periods, and different risk-measures), second, subgroups relating to competing ways to gain data input (i.e., different education definitions, weighted versus unweighted samples, different types of clustering, different time period from which income parameters are estimated), third, subgroups relating to competing L_{p}-norms used to estimate risk-attitude, and fourth, subgroups relating to competing transformations rules to code the binary dependent variables.

4.3.2 Gender and risk-attitudes

4.3.2.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

In my study the relation between gender and the probability of being risk-averse is captured by the dummy variable Male. The relation of Male to the probability of being risk-averse can be statistically investigated for about 99% of the 4,536 ways to estimate risk-attitude, i.e., (quasi-) complete separation is not a problem. Based on this sample, Male does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Concerning significance, no statement can be made since Leamer's criterion of robustness does not find a robust relation of Male to any of the dependent variables. Yet, I derive a strong tendency for a positive coefficient of Male which means men might have a higher probability of being risk-averse than women.

Comparison to the literature and interpretation

Both non-robustness and the strong tendency for a positive sign contradict the consensus in the literature where men are found to be significantly more willing to take risks than women in non-migration domains.[288] To the best of my knowledge only Harrison, Lau, and Rutström (2007) do not find significant differences in risk-taking between men and women when investigating answers to lottery questions in Denmark.

This difference to findings of the literature is not surprising to me for two reasons. First, risk-attitudes are domain-specific and the domain of migration has not been surveyed in the literature yet. The fact that women are usually more risk-averse than men when taking financial decisions or health risks – as investigated by other authors - does not necessarily mean that this also holds for the migration decision. Second, risk-attitudes could truly be related to certain personal traits rather than gender, while these very traits could be related with male in the population. Since I survey a certain subgroup of the population, namely migrants, it could be that the generally found relation between certain traits and gender cannot be found in my subgroup of migrants.

[288] For a detailed discussion of findings on the relation for gender and risk-attitudes in the previous empirical literature please refer to Part C, p. 168.

4.3.2.2 Detailed statistical reasoning

All statistical results on the influence of Male on the probability of being risk-averse are summarized in Table 37, 201.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables n	missing values % of 4,536	Leamer significant % of n	positive % of significant		Sala-I-Martin weighted CDF(0) ≥95% % of n	positive % of robust		positive % of n		Sala-I-Martin unweighted CDF(0) ≥95% % of n	positive % of robust		positive % of n	
Total	4,478	1%	0%	n.a.		19%	73%		76%	(+)	12%	64%		77%	(+)
Ind	3,840	1%	0%	n.a.		20%	74%		77%	(+)	12%	68%		78%	(+)
Fam	638	2%	0%	n.a.		17%	65%		72%		15%	40%		71%	
Ann	1,492	1%	0%	n.a.		20%	72%		76%	(+)	13%	64%		77%	(+)
Wor	1,494	1%	0%	n.a.		19%	74%		76%	(+)	12%	64%		77%	(+)
Lif	1,492	1%	0%	n.a.		19%	73%		76%	(+)	12%	63%		77%	(+)
Var	2,268	0%	0%	n.a.		13%	90%	(+)	83%	(+)	7%	99%	(+)	86%	(+)
LP2	2,210	3%	0%	n.a.		26%	64%		70%		18%	49%		68%	
Ed1	1,274	2%	0%	n.a.		19%	68%		77%	(+)	14%	52%		75%	(+)
Ed2	636	2%	0%	n.a.		21%	71%		79%	(+)	12%	61%		78%	(+)
Ed3	648	0%	0%	n.a.		14%	61%		66%		9%	45%		66%	
Ed4	648	0%	0%	n.a.		13%	55%		59%		9%	46%		62%	
Ed5	636	2%	0%	n.a.		27%	86%	(+)	88%	(+)	14%	90%	(+)	93%	(+)
Ed6	636	2%	0%	n.a.		24%	85%	(+)	88%	(+)	13%	89%	(+)	91%	(+)
Wei	2,220	2%	0%	n.a.		20%	74%		78%	(+)	12%	63%		82%	(+)
Unw	2,258	0%	0%	n.a.		19%	72%		75%	(+)	13%	65%		72%	
Sep	2,258	0%	0%	n.a.		18%	61%		70%		12%	40%		71%	
Poo	2,220	2%	0%	n.a.		21%	83%	(+)	83%	(+)	13%	87%	(+)	83%	(+)
One	1,512	0%	0%	n.a.		23%	76%	(+)	78%	(+)	12%	95%	(+)	89%	(+)
Ad1	1,483	2%	0%	n.a.		16%	67%		74%		10%	38%		70%	
Ad2	1,483	2%	0%	n.a.		19%	75%	(+)	76%	(+)	14%	55%		72%	
L1	1,512	0%	0%	n.a.		14%	87%	(+)	84%	(+)	8%	68%		82%	(+)
L2	1,512	0%	0%	n.a.		19%	66%		71%		14%	36%		70%	
Ma	1,454	4%	0%	n.a.		25%	70%		74%		15%	88%	(+)	79%	(+)
2Kat3	1,512	0%	0%	n.a.		17%	90%	(+)	85%	(+)	8%	81%	(+)	82%	(+)
2Kat1	1,484	2%	0%	n.a.		15%	90%	(+)	78%	(+)	10%	67%		78%	(+)
4Kat2	1,482	2%	0%	n.a.		26%	52%		66%		19%	54%		71%	

Table 37: Results of the extreme bounds analysis for the independent variable Male aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Sample size and (quasi-) complete separation

Since quasi-complete separation only occurs for 1% of the 4,536 dependent variables (Column (3), Table 37, p. 201) the statistical reasoning on robustness, sign, and significance of Male is based on a sufficiently high sample of 4,478 dependent variables (Column (2), Table 37, p. 201).

Concerning the 58 dependent variables (about 1% of 4,536, Column (3), Table 37, p. 201) where quasi-complete separation occurs, it is noteworthy that all men in the sample are risk-averse.[289] In contrast, women can be found in the group of risk-averse as well as risk-seeking migrants.

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, Male is clearly classified as fragile, i.e., Males does not show a robust influence on any of the 4,536 dependent variables (Column (4), Table 37, p. 201). Since no robust relation is found, a statement on significance cannot be derived (Columns (5) and (6), Table 37, p. 201).

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, Male is also clearly classified as fragile. Sala-I-Martin's weighted criterion finds only 19% of the 4,478 dependent variables to be robustly related to the probability of being risk-averse (Column (7), Table 37, p. 201)[290], and Sala-I-Martin's unweighted criterion finds only 12% of the 4,478 dependent variables to be robustly related to the probability of being risk-averse (Column (12), Table 37, p. 201).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, a strong tendency for a positive sign of the coefficient of Male is observable since all percentages available guide in the same direction and two of them meet the 75%-criterion (indicated by (+) in Columns (11) and (16), Table 37, p. 201). In detail this means: The percentage of positive relations found by Leamer's criterion cannot be considered because Leamer's criterion does not find any robust relations. Among the 19% robust relations by Sala-I-Martin's weighted criterion, 73% find a robust positive influence of Male, while the remaining 27% find a robust negative influence (Column (8), Table 37, p. 201). If percentages are not restricted to robust relations, Sala-I-Martin's weighted criterion estimates a greater probability for a positive sign of the

[289] Note that for each of the 58 dependent variables for which quasi-separation occurs, there are 264 men in my sample accounting for 82% of all migrants. All of these 264 men are risk-averse as measured by the respective dependent variable.

[290] In detail, the 19% mean that for 19% of the dependent variables for which an extreme bounds analysis could have been run, the probability for either a positive or a negative sign of the coefficient is 95% or higher.

coefficient of Male for 76% of the 4,478 dependent variables statistically investigated (Column (10), Table 37, p. 201). The situation for Sala-I-Martin's unweighted criterion is alike. Among the 12% robust relations, 64% find a robust positive influence of Male (Column (13), Table 37, p. 201), while among all 4,478 dependent variables, 77% find a greater probability for a positive sign (Column (15), Table 37, p. 201).

4.3.2.3 Sensitivity for subgroups of dependent variables

4.3.2.3.1 Overview

Robustness

The fragility and non-significance of Male in explaining the probability of being risk-averse over the 4,478 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, one finding is noteworthy: Comparing competing risk-measures, I find that the subgroup referring to risk-measure semi-variance (LP2) exhibits about two times the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria as the subgroup referring to risk-measure variance (Var). This indicates that the relation between Male and the probability of being risk-averse is stronger for the downside-risk (semi-variance) than for risk that includes any variation of income (variance).

Tendency for a sign

The finding of a strong tendency for a positive sign of Male for the 4,478 dependent variables statistically investigated is not sensitive to (i) variations of the planning period (denoted by Ann, Wor, and Lif), (ii) whether missing education information is brought forward from years before or after the move (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), and not sensitive to (iv) whether income parameters are estimated from annual income data of one year or three years of data - as long as data is not adjusted for inflation (denoted by One and Ad1).

However, the strong tendency for a positive sign is sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) different risk-measures (denoted by Var and LP2), (iii) different education definitions (denoted by Ed1 to Ed6), (iv) different types of clustering (denoted by Sep and Poo), (v) whether three year's data from which income parameters are estimated is adjusted for inflation or not (denoted by Ad1 and Ad2), (vi) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (vi) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.2.3.2 Detailed statistical reasoning and interpretation

One fact is identical in all subgroups and will therefore not be mentioned again for every subgroup. That is, the relation between Male and the probability of being risk-averse by Leamer's criterion is not robust and therefore not significant for any subgroup. Consequently, robustness and the tendency for a sign are only discussed based on Sala-I-Martin's weighted and unweighted criteria.

<u>Detailed comparison of subgroups of dependent variables relating to different decision problems</u>

- The relation of Male to the probability of being risk-averse is sensitive to whether the migration decision is estimated based on **Head's individual income or family income** (denoted by Ind and Fam). First, it is noteworthy that the percentages of dependent variable for which quasi-complete separation occurs are similar in both subgroups although the number of dependent variables that refer to Head's individual income and family income differ considerably (3,840 compared to only 638, Column (2), Table 37, p. 201). Second, Male is fragile in both subgroups. Third, the subgroup referring to Head's individual income (Ind) exhibits a strong tendency for a positive sign of the coefficient of Male. In contrast, the subgroup referring to family income (Fam) exhibits no tendency for a sign of the coefficient of Male.

  I interpret these findings as follows: First, it seems as if male Heads have a higher probability of being risk-averse concerning variations of their own income than female Heads. In contrast, for the probability of being risk-averse concerning variations of family income, no difference between male and female Heads can be detected. Put differently, while men and women are equally worried about variations of family income, male Heads are more anxious about variations of their own income than female Heads. Second, the findings for the two subgroups can also be due to the considerably lower number of dependent variables investigated by the subgroup referring to family income, which amount to only one sixth of the number of dependent variables investigates by the subgroup referring to Head's individual income.

- The relation of Male to the probability of being risk-averse is not sensitive to **different planning periods** investigated in my study, i.e., one year, time until reaching full retirement age, and time until reaching life expectancy (denoted by Ann, Wor, and Lif). In detail, I find that, first, percentages of quasi-complete separation are equally low. Second, Male is fragile in both subgroups. Third, in all subgroups a strong tendency for a positive sign is found.

I interpret these findings as follows: The non-sensitivity of the results on Male for different planning periods might be due to the way annual income parameters have been used to calculate income parameters for longer planning periods (see Part A, Section 2.4.2 for a detailed discussion). Recall that it was assumed that annual nominal income is constant; however, the length over which this income is paid varies for different planning periods. This means, although differences in the relation of annual mean and variance of income (or semi-variance of income, respectively) between U.S. states are even more pronounced for longer planning periods, the rank of U.S. states does not change when the planning period is modified. Therefore, estimations of risk-attitudes might be similar.

- The relation of Male to the probability of being risk-averse is sensitive to **different risk-measures** variance and semi-variance (denoted by Var and LP2). First, percentages of quasi-complete separation are similar. Second, although Male is classified as fragile in both subgroups, applying semi-variance (LP2) results in about twice as much robust relations by Sala-I-Martin's weighted and unweighted criteria compared to when risk-measure variance (Var) is applied. Third, concerning the tendency for a sign, results are very sensitive to different risk-measures. The subgroup referring to variance (Var) shows a strong tendency for a positive sign. In contrast, the subgroup referring to semi-variance (LP2) shows no tendency for a sign.

  I interpret these findings as follows: Results indicate that male Heads compared to female Heads have a higher probability of being risk-averse towards any variation of income (Var), but their probability of being risk-averse towards risk that includes only scenarios with a shortfall below the expected income (LP2) has no tendency to be higher or lower compared to female Heads. Put differently, while men and women are equally worried about downside-risk, it is male Heads that are more anxious about any variation of income than female Heads. Note that the different size of risk-attitude parameters for variance and semi-variance (see Part C, Section 1.4 p. 152) cannot explain the observed sensitivity since risk-attitudes were transformed to binary variables.

Detailed comparison of subgroups of dependent variables relating to different ways to gain data input

- The relation of Male to the probability of being risk-averse is sensitive to **different education definitions** (denoted by Ed1 to Ed6). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income only education definition one was applied. Therefore, the number of dependent variables to be analyzed amounts to 648 for education definitions two to six, while education definition one additionally includes all migration decisions based on family income.

Comparing the results for the six subgroups referring to the six different education definitions, I find that although the results for the six subgroups are not similar over all subgroups, the results are pair wise similar as follows:

- The results for subgroups referring to education definitions one and two (denoted by Ed1 and Ed2) are similar. First, for both subgroups quasi-complete separation occurs for 2% of the 4,536 dependent variables (Column (3), Table 37, p. 201). Second, Male is classified as fragile in both subgroups. Third, both subgroups show a strong tendency for a positive sign.

- The results for subgroups referring to education definitions three and four (denoted by Ed3 and Ed4) are similar. First, for both subgroups quasi-complete separation does not occur. Second, Male is classified as fragile in both subgroups. Third, both subgroups show no tendency for a sign.

- The results for subgroups referring to education definitions five and six (denoted by Ed5 and Ed6) are similar. First, for both subgroups quasi-complete separation does not occur. Second, Male is classified as fragile in both subgroups. Third, both subgroups exhibit a strong tendency for a positive sign where all four percentages referring to robust/all relations, applying Sala-I-Martin's weighted/unweighted criterion are greater than 75%. Compared to all other subgroups (Ed1 to Ed4), the positive tendency is most pronounced for subgroups Ed5 and Ed6.

To be able to interpret the findings of pairwise similar subgroups, recall that the six education definitions were derived from two competing solutions to each of the two following questions: (1) The question on how to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File, and (2) whether education variables from years before or after the move should be considered first when education variables are not available in each year.

Therefore, I interpret these findings of the sensitivity to different education definitions as follows: The pairs of subgroups showing similar results only differ in their answer to Question (2). Subgroups Ed1, Ed3, and Ed5 first search for education information in the years centered around the moving date, while subgroups Ed2, Ed4, and Ed6 first search in the years before the move and then in the years after the move. Obviously, the different answers to Question (2) do not alter the result of how Male is related to the probability of being risk-averse in the migration context. In contrast, it makes a difference how Question (1) is answered, i.e., whether

(i) education information from the Single-Year Family Files is always preferred in each year over data from the Cross-Year Individual File (subgroups Ed1, Ed2, strong tendency for a positive sign), or whether (ii) data from the Single-Year Family Files of all years is first considered before data from the Cross-Year Individual File is considered (subgroups Ed3, Ed4, no tendency for a sign), or whether (iii) it is simply the highest education level of both the Single-Year Family Files and the Cross-Year Individual File that is preferred in each year (subgroups Ed 5, Ed6, most pronounced strong tendency for a positive sign). The difference in the methodology of answers (i) and (iii) seems to be qualitative marginal since the related subgroups Ed1, Ed2, Ed5, and Ed6 all show a strong tendency for a positive sign. I interpret these findings in the way that education information reported in the Single-Year Family Files is most often higher than education information in the Cross-Year Individual File. The difference in the methodology of answer (ii) is responsible for the sensitivity of Male in explaining the probability of being risk-averse for different education definitions.

- The relation of Male to the probability of being risk-averse is not sensitive to whether **income parameters are estimated based on weighted or unweighted samples** (denoted by Wei and Unw). For both subgroups I find that, first, percentages of quasi-complete separation are equally low, second, Male is classified as fragile, and third, a strong tendency for a positive sign is found.

  I interpret these findings as follows: It does not make a difference for the results of Male whether the sample from which income parameters are estimated is unweighted or weighted to be representative of the population. The fear was that when sample sizes for certain socio-economic groups are critically small – this is the case for some socio-economic groups (see Part B, Section 3.5) – weighting the sample would bias the estimation of income parameters, and therefore, bias the estimation of risk-attitudes. Obviously, this concern cannot be confirmed for the relation of Male to the probability of being risk-averse in my study.

- The relation of Male to the probability of being risk-averse is sensitive to the **type of clustering** of people from which income parameters are estimated, i.e., separate clustering for each year versus pooled clustering with same clusters in all years (denoted by Sep and Poo). First, percentages of quasi-complete separation are equally low in both subgroups. Second, Male is fragile in both subgroups. Third, subgroups show opposing tendencies for a sign, i.e., the subgroup referring to separate clustering shows no tendency, while the subgroup referring to pooled clustering shows a strong tendency for a positive sign.

I interpret these findings as follows: The clustering was performed along age and education to reduce the number of socio-economic groups from which income parameters were estimated in order to reach higher sample sizes in the corresponding socio-economic groups. The two types of clustering resulted in the same education-clusters, but in considerably different age-clusters (i.e., 20 year old men were clustered together with all men between 16 and 20 years of age following one type of clustering; the same person was clustered together with all men between 19 and 25 year of age following the other type of clustering). Hence, it is not surprising that estimated income parameters and risk-attitudes also differ considerably.

- The relation of Male to the probability of being risk-averse is sensitive to different **time periods from which income parameters are estimated**, namely annual income data of one single year, annual income date from three years around the moving date adjusted for inflation, and annual income data from three years around the moving date not adjusted for inflation (denoted by One, Ad1, and Ad2). For all subgroups, I find that, first, percentages of quasi-complete separation are low and, second, Male is fragile. Third, concerning the tendency for a sign, I discover a strong tendency for a positive sign for the subgroup referring to three year's data without inflation adjustment (Ad2) and the subgroup referring to one year's data (One), where this tendency is even more pronounced for the latter (One). The subgroup referring to three year's data with inflation adjustment (Ad1) shows no tendency for a sign.

  I interpret these findings as follows: Results are not sensitive to whether income parameters are estimated from one year's data (One) or three year's data as long as the three year's data is not adjusted for inflation (Ad2). Although inflation adjustment seems reasonable in general, it is explicitly not recommended by the editor of the American Community Survey.[291] The reason is that respondents of the American Community Survey are asked to report their income in the last 12 month. Since interviews are run throughout the year, income data of neighboring years of data already refer to overlapping reference periods. Inflation adjustment would therefore bias results in an uncontrollable way. In my study, this warning of the editor of the American Community Survey proves to be true for the coefficient of Male.

Detailed comparison of subgroups of dependent variables relating to different estimation procedures of risk-attitudes

The relation of Male to the probability of being risk-averse is sensitive to **different measurements of predictive errors**, namely L_p-norms one-, two- and infinity-norms where $p = \{1,2,\infty\}$ (denoted by

[291] See Minnesota Population Center, University of Minnesota [No date b].

L1, L2, and Ma). First, the percentages of quasi-complete separation are equally low in all subgroups. Second, although Male is fragile in all subgroups, the percentages of robust relations are highest when the maximum of predictive errors is minimized (subgroup Ma with 25% and 15% by Sala-I-Martin's weighted and unweighted criterion, respectively (see Columns (7) and (12), Table 37, p. 201)) and lowest when the sum of predictive errors is minimized (subgroup L1 with 14% and 8% by Sala-I-Martin's weighted and unweighted criterion, respectively (see Columns (7) and (12), Table 37, p. 201)). Third, subgroups show different tendencies for a sign. The subgroup referring to minimizing the sum of predictive errors (L1) and the subgroup referring to minimizing the maximum of predictive errors (Ma) show a strong tendency for a positive sign. In contrast, the subgroup referring to minimizing the squared predictive error shows no tendency for a certain sign (L2).

I interpret these findings as follows: Differences between male and female Heads are most pronounced when greater weight is put to greater predictive errors (L2), while results for the subgroups referring to minimizing the sum of predictive errors (L1) and minimizing the maximum of predictive errors (Ma) are qualitative not different. This fact indicates that cases with higher but not maximal predictive errors are responsible for the different results of Male for different measurements of predictive errors.

<u>Detailed comparison of subgroups of dependent variables relating to different transformation rules to code the binary dependent variables</u>

The relation of Male to the probability of being risk-averse is sensitive to **different transformation rules** (denoted by 2Kat1, 2Kat3, and 4Kat2). First, percentages of quasi-complete separation are equally low in all subgroups. Second, although Male is classified as fragile in all subgroups, I find the greatest percentage of robust relations for the subgroup including only the 50% migrants exhibiting the most pronounced degree of risk-attitude (4Kat2). Third, a strong tendency for a sign can be found in the subgroup referring to the transformation rule that includes all migrants (2Kat3) and the subgroup referring to the transformation rule that includes only the 95% migrants exhibiting the strongest degree of risk-attitude (2Kat1), while the subgroup referring to the most extreme transformation rule (4Kat2) only indicates a vague tendency for a positive sign.

I interpret these findings as follows: Among the migrants that exhibit a lower degree of risk-attitude (2Kat3 and 2Kat2), male and female Heads are more different in their risk-attitudes than migrants that exhibit a strong degree of risk-attitude (i.e., extremely risk-averse or extremely risk-seeking (4Kat2)).

4.3.3 Age and risk-attitudes

4.3.3.1 Overview of results irrespective of the way risk-attitudes are estimated

<u>**Significance, robustness, and tendency for a sign**</u>

The relation of Age to the probability of being risk-averse can be statistically investigated for all 4,536 dependent variables since (quasi-) complete separation does not occur. Based on this sample, Age does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Concerning significance, no statement can be made since Leamer's criterion of robustness does not find a robust relation of Age to any of the dependent variables.

A tendency for a sign cannot be derived. Instead the following pattern can be observed: The majority of robust relations by Sala-I-Martin's weighted and unweighted criteria finds a negative sign of Age, while the majority of all relations irrespective of their robustness finds a positive sign of Age. Obviously, it does not make a difference whether Sala-I-Martin's criterion is applied in its weighted or unweighted version, but whether the analysis is restricted to robust relations or not.

<u>**Comparison to the literature and interpretation**</u>

Both non-robustness and no tendency for a sign of the relation of Age to the probability of being risk-averse contradict the consensus in the literature, where Age has been found to be significantly positive related to risk-aversion in the non-migration domain by almost all studies.[292] To the best of my knowledge, only Wik, Kebede, Bergland, and Holden (2004) do not find a significant effect of age among households in Northern Zambia using an experimental gambling approach.

There are three explanations for my findings that the age effect is not robust and has no tendency for a sign: First, risk-attitudes are domain-specific, which means the findings of the previous literature relating to non-migration domains do not necessarily have to be true in the context of migration. Second, I investigate only a subgroup of the population, namely migrants, that are not necessarily similar to the population. That the relation of age on risk-attitudes may be different for different subgroups of the population has already been shown by Hartog, Ferrer-i-Carbonell, and Jonker (2002). It is noteworthy, that Hartog, Ferrer-i-Carbonell, and Jonker (2002) can confirm the findings of the previous literature for the data that can be expected to be more similar to the population (i.e., the data set of the 25,000 readers of several Dutch newspapers). In contrast, they find the opposite

[292] For a detailed discussion of findings on the relation for age and risk-attitudes in the previous empirical literature please refer to Part C, p. 169.

effect for the more homogenous group of accountants. It is therefore not surprising that I cannot confirm previous findings for the subgroups of migrants. The third argument might explain why no tendency for a sign of the coefficient is found as follows: As already discussed in Part C, p. 169, it is possible that an additional nonlinear age-effect on the probability of being risk-averse exists. If a nonlinear age-effect exists, but the relation is captured only by a linear variable like Age, the sign of the linear coefficient crucially depends on the location of the extremum of the age effect. For example, if the probability of being risk-averse is inverse U-shaped (i.e., rises with age until a certain maximum is reached and falls thereafter), the estimated linear regression coefficient will be the more positive the greater the age at which the maximum is reached. Generalizing this insight to my study this means, if a non-linear age-effect exists and its estimated extremum is located at different ages for different dependent variables, it is not surprising that the regression coefficient of the linear age-effects switches its sign for different depended variables. Whether a non-linear age-effect exists, will be discussed in Part C, Section 4.3.5, where the empirical results on AgeSquared are interpreted.

4.3.3.2 Detailed statistical reasoning

All statistical results on the influence of Age on the probability of being risk-averse are summarized in Table 38, p. 212.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,536	0%	0%	n.a.		17%	21%	(-)	60%		3%	32%		60%	
Ind	3,888	0%	0%	n.a.		16%	22%	(-)	60%		3%	31%		61%	
Fam	648	0%	0%	n.a.		19%	19%	(-)	56%		5%	34%		55%	
Ann	1,512	0%	0%	n.a.		16%	23%	(-)	60%		4%	32%		60%	
Wor	1,512	0%	0%	n.a.		18%	21%	(-)	59%		4%	35%		60%	
Lif	1,512	0%	0%	n.a.		17%	20%	(-)	59%		3%	29%		60%	
Var	2,268	0%	0%	n.a.		11%	54%		82%	(+)	2%	100%	(+)	89%	(+)
LP2	2,268	0%	0%	n.a.		22%	4%	(-)	37%		5%	0%	(-)	31%	
Ed1	1,296	0%	0%	n.a.		18%	18%	(-)	56%		4%	31%		54%	
Ed2	648	0%	0%	n.a.		17%	31%		56%		4%	48%		55%	
Ed3	648	0%	0%	n.a.		15%	22%	(-)	70%		3%	0%	(-)	74%	
Ed4	648	0%	0%	n.a.		12%	5%	(-)	65%		2%	0%	(-)	72%	
Ed5	648	0%	0%	n.a.		18%	17%	(-)	58%		3%	29%		56%	
Ed6	648	0%	0%	n.a.		19%	32%		57%		4%	61%		54%	
Wei	2,268	0%	0%	n.a.		14%	2%	(-)	56%		2%	0%	(-)	58%	
Unw	2,268	0%	0%	n.a.		19%	35%		63%		5%	48%		62%	
Sep	2,268	0%	0%	n.a.		15%	15%	(-)	60%		4%	31%		62%	
Poo	2,268	0%	0%	n.a.		19%	26%		59%		3%	33%		58%	
One	1,512	0%	0%	n.a.		13%	16%	(-)	44%		3%	35%		45%	
Ad1	1,512	0%	0%	n.a.		19%	26%		67%		4%	45%		68%	
Ad2	1,512	0%	0%	n.a.		18%	20%	(-)	68%		3%	13%	(-)	68%	
L1	1,512	0%	0%	n.a.		9%	20%	(-)	58%		1%	0%	(-)	62%	
L2	1,512	0%	0%	n.a.		13%	27%		66%		3%	0%	(-)	60%	
Ma	1,512	0%	0%	n.a.		28%	19%	(-)	55%		6%	53%		58%	
2Kat3	1,512	0%	0%	n.a.		10%	41%		72%		0%	100%	(+)	69%	
2Kat1	1,512	0%	0%	n.a.		13%	24%	(-)	67%		1%	14%	(-)	64%	
4Kat2	1,512	0%	0%	n.a.		28%	12%	(-)	41%		9%	32%		47%	

Table 38: Results of the extreme bounds analysis for the independent variable Age aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Sample size and (quasi-) complete separation

(Quasi-) Complete separation does not occur. Hence, the statistical reasoning on robustness, sign, and significance of Age is based on all 4,536 dependent variables (Column (2), Table 38, p. 212).

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, Age is clearly classified as fragile, i.e., Age does not show a robust influence on any of the 4,536 dependent variables (Column (4), Table 38, p. 212). Since no robust relation is found, a statement on significance cannot be derived (Columns (5) and (6), Table 38, p. 212).

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, Age is also clearly classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 17% (3%) of the 4,536 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 38, p. 212).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, no tendency for a sign of the coefficient of Age is observable. Note that the percentage of positive relations found by Leamer's criterion cannot be considered because Leamer's criterion does not find any robust relations.

4.3.3.3 Sensitivity for subgroups of dependent variables

4.3.3.3.1 Overview

Robustness

The fragility and non-significance of Age in explaining the probability of being risk-averse over all 4,536 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing results of subgroups referring to different risk-measures, I find that the subgroup referring to semi-variance (LP2) exhibits about two times the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria as the subgroup referring to variance (Var). Second, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are

deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of no tendency for a sign of Age for all 4,536 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the education definition (more precisely, results are neither sensitive to competing answers on whether education variables from years before or after the move should be preferred, nor sensitive to competing answers on how to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File; related subgroups are denoted by Ed1 to Ed6), (iv) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (v) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (vi) whether income parameters are estimated from annual income data of three years of data that is adjusted for inflation or not (denoted by Ad1 and Ad2), and not sensitive to (vii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma).

However, the finding of no tendency for a sign is sensitive to (i) different risk-measures (denoted by Var and LP2), (ii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), and sensitive to (iii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.3.3.2 Detailed statistical reasoning and interpretation

Two facts are identical for all subgroups and will therefore not be mentioned again for every subgroup. They are first, (quasi-) complete separation does not occur for the total of all 4,536 dependent variables and is therefore also of no concern for any subgroup of dependent variables. Second, the relation between Age and the probability of being risk-averse is not robust for any subgroup by Leamer's criterion of robustness. Consequently, only robustness by Sala-I-Martin's criterion and related tendencies for a sign are discussed in this section.

Detailed comparison of subgroups of dependent variables relating to different model specifications

- The relation of Age to the probability of being risk-averse is not sensitive to whether the migration decision is estimated based on **Head's individual income or family income** (denoted by Ind and Fam). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income only education definition one was applied. Therefore, the number of dependent variables in the subgroup referring to family income (Fam) amounts to one sixth of the dependent variables based in Head's individual income (Ind). For both subgroups I find that, first, Age is classified as fragile, and second, no tendency for a sign can be observed.

I interpret these findings as follows: Irrespective of whether Head's individual income or family income is considered, it is not clear whether the probability of being risk-averse increases or decreases when people turn one year older.

- The relation of Age to the probability of being risk-averse is not sensitive to **different planning periods** investigated in my study, i.e., one year, time until reaching full retirement age, and time until reaching life expectancy (denoted by Ann, Wor, and Lif). For all subgroups I find that, first, Age is classified as fragile, and second, no tendency for a sign can be observed.

  I interpret these findings as being due to the way income parameters for longer planning periods have been estimated (for a detailed discussion of this argument see the corresponding interpretation for Male, p. 204).

- The relation of Age to the probability of being risk-averse is sensitive to **different risk-measures** variance and semi-variance (denoted by Var and LP2). First, although Age is classified as fragile in both subgroups, applying semi-variance (LP2) results in about twice as much robust relations by Sala-I-Martin's weighted and unweighted criteria compared to risk-measure variance (Var). Second, concerning the tendency for a sign of the coefficient of Age, I find the complete opposite for both subgroups. The subgroup referring to risk-measure variance (Var) shows a strong tendency for a positive sign, while the subgroup referring to semi-variance (LP2) shows a strong tendency for a negative sign.

  I interpreted the findings concerning robustness as follows: First, it could be that the relation between age and the probability of being risk-averse towards the downside-risk (LP2) is approximately linear, while the same relation for risk-measure variance (Var) is non-linear. Second, if both relations are non-linear, the higher percentage of robust relations in the subgroup referring to risk-measure semi-variance (LP2) indicates that the probability of being risk-averse can better be approximated by a linear function for risk-measure semi-variance (LP2) compared to variance (Var). Third, if both relations were approximately linear, the considerably lower percentage of robust relations in the subgroup referring to risk-measure variance (Var) indicates that people of different ages are more similar in their risk-attitude towards variance than semi-variance.

I interpreted the findings concerning the tendency for a sign as follows: For older migrants the probability of being risk-averse towards the downside-risk decreases, while the probability for being risk-averse towards any variation increases. This can be explained if older migrants consider deviations of income above the expected value as more problematic compared to younger migrants than deviations below the expected value. A possible explanation for this finding might be that older migrants have a better downside protection compared to younger migrants, e.g., older migrants own assets that are negatively correlated with income. In this case, their total wealth might decreases when income increases.

Detailed comparison of subgroups of dependent variables relating to different ways to gain data input

- The relation of Age to the probability of being risk-averse is not sensitive to **different education definitions** (denoted by Ed1 to Ed6). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income only education definition one was applied. Therefore, the number of dependent variables to be analyzed amounts to 648 for education definitions two to six, while education definition one additionally includes all migration decisions based on family income. For all six subgroups I find that, first, Age is classified as fragile, and second, no tendency for a sign can be observed.

  I interpret these findings as follows: Since the six education definitions are due to competing answers on (1) whether education variables from years before or after the move should be preferred, and (2) how to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File, I conclude that the relation of Age to the probability of being risk-averse is not sensitive to competing answers on both questions.

- The relation of Age to the probability of being risk-averse is not sensitive to whether **income parameters are estimated based on weighted or unweighted samples** (denoted by Wei and Unw). For both subgroups I find that, first, Age is classified as fragile, and second, no tendency for a sign can be observed.

I interpret these findings as follows: Obviously, the concern of biased results due to biased estimations of the income parameters for weighted samples cannot be confirmed for the relation of Age to the probability of being risk-averse in my study.

- The relation of Age to the probability of being risk-averse is not sensitive to the **type of clustering** of people from which income parameters are estimated, i.e., separate clustering for each year versus pooled clustering with same clusters in all years (denoted by Sep and Poo). For both subgroups I find that, first, Age is classified as fragile, and second, no tendency for a sign can be observed.

  I interpret these findings as follows: Although the age-clusters resulting from the two types of clustering differed considerably, it does not make a difference for the effect of Age on the probability of being risk-averse whether age-clusters are defined separately for each year (Sep) or whether age-clusters are defined in the same way over all years (Poo).

- The relation of Age to the probability of being risk-averse is sensitive to different **time periods from which income parameters are estimated**, namely annual income data of one single year, annual income date from three years around the moving date adjusted for inflation, and annual income data from three years around the moving date not adjusted for inflation (denoted by One, Ad1, and Ad2). First, Age is classified as fragile in both subgroups. Second, I find a strong tendency for a negative sign in the subgroup referring to one year's data (One), while no tendency for a sign can be derived for the two subgroups referring to three year's data (Ad1 and Ad2).

  I interpret these findings as follows: It seems as if the results of Age and its relation to the probability of being risk-averse are not sensitive to whether three year's data is adjusted for inflation or not (Ad1 versus Ad2). This is in contrast to the concern expressed by the editor of the American Community Survey[293] (for a detailed discussion of this argument see the corresponding interpretation for Male, p. 208). However, results of Age and its relation to the probability of being risk-averse are sensitive whether income parameters are estimated from one year's data versus three year's date (One versus Ad1, Ad2).

[293] See Minnesota Population Center, University of Minnesota [No date b].

<u>Detailed comparison of subgroups of dependent variables relating to different estimation procedures of risk-attitudes</u>

The relation of Age to the probability of being risk-averse is not sensitive to **different measurements of predictive errors**, namely L_p-norms one-, two- and infinity-norms where $p = \{1,2,\infty\}$ (denoted by L1, L2, and Ma). First, although Age is classified as fragile in all subgroups, the percentage of robust relations by Sala-I-Martin's weighted and unweighted criteria is highest when the maximum of predictive errors is minimized (Ma) and lowest when the sum of predictive errors is minimized (L1). Second, no tendency for a sign of Age can be derived for any subgroup

I interpret these findings as follows: It seems as if the maximum predictive error is decisive since results for subgroups that additionally account for smaller predictive errors (L1 and L2) exhibit qualitative the same results. This argument might also explain why the subgroup referring to minimizing the maximum error is also the subgroup with the highest percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria.

<u>Detailed comparison of subgroups of dependent variables relating to different transformation rules to code the binary dependent variables</u>

The relation of Age to the probability of being risk-averse is sensitive to **different transformation rules** (denoted by 2Kat1, 2Kat3, and 4Kat2). First, although Age is classified as fragile in all subgroups, I find the smallest percentage of robust relations for the subgroup that includes all migrants (2Kat3), followed by the subgroup that includes only the 95% migrants exhibiting the strongest degree of risk-attitude (2Kat1), and the greatest percentage of robust relations for the subgroup including only the 50% migrants exhibiting the strongest degree of risk-attitude (4Kat2). Second, only for the subgroup referring to the 50% migrants that exhibits the most extreme degree of risk-attitude (4Kat2) a strong tendency for a negative sign can be observed. For the other subgroups no tendency for a sign is observed.

I interpret these findings as follows: The idea of the extreme transformation rule (4Kat2) was that once migrants with less pronounced degrees of risk-attitudes are deleted from the sample, the true relation between an independent variable and the probability of being risk-averse would be more pronounced. This is exactly what can be seen for the independent variable Age, i.e., for the subgroup including only migrants with the most extreme degree of risk-attitude (4Kat2): The percentage of robust relations is greatest and it is the only subgroup for which a tendency for a sign can be observed.

4.3.4 Education and risk-attitudes

4.3.4.1 Reparatory remarks

Central methodology of interpretation for education (captured by Edu2 and Edu3)

In my study education is captured by the two dummy variables Edu2 and Edu3. To derive a statement on the relation of education on the probability of being risk-averse, both variables have to be jointly interpreted. This is possible because Edu2 and Edu3 are fixed variables that are simultaneously included in all regressions. In detail this means the central methodology of interpretation is applied as follows:

- Statements on **sample size and (quasi-) complete separation** are identical for both dummy variables since results of both dummy variables are based on identical regressions.

- To judge the **joint significance of Edu2 and Edu3** on the probability of being risk-averse, I run additional Wald-tests on joint significance of both dummy variables Edu2 and Edu3 for all over 12 million regressions. To finally derive a statement on joint significance of Edu2 and Edu3 over all 4,536 dependent variables, I aggregate results in three steps analogue to the central methodology applied for individual independent variables as follows: First, for each dependent variable, I judge Edu2 and Edu3 to be jointly significant (at some significance level p) if both dummy variables are jointly significant for at least 95% of the regressions run on that dependent variable. Second, to derive a statement on the joint significance of Edu2 and Edu3 (at some significance level p), irrespective of the way risk-attitude is estimated, I aggregate results for all 4,485 dependent variables by calculating the percentage of significant relations (at some significance level p) found among all dependent variables analyzed. Third, in analogy to the 75%-criterion applied to individual variables, I require Edu2 and Edu3 to be jointly significant for at least 75% of the dependent variables in order to classify education as significant in explaining the probability of being risk-averse.

- **Robustness by Leamer's and Sala-I-Martin's criteria** are checked separately for Edu2 and Edu3 where education is classified as robust only if both Edu2 and Edu3 are classified as robust.

- **The tendency for a sign** is derived for both dummy variables separately but can be jointly interpreted. This joint interpretation is possible since Edu2 and Edu3 both have the same reference category (i.e., Edu1), where Edu1 indicates the lowest, Edu2 the next higher, and Edu3 the highest level of education.

4.3.4.2 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

I find that the relation of Head's education to the probability of being risk-averse can be statistically investigated for about 99% of the 4,536 ways to estimate risk-attitude, i.e., (quasi-) complete separation is not a problem. Based on this sample, Edu2 and Edu3 are jointly significant at a 1% significance level for 36% of the 4,485 dependent variables (Column (6), Table 39, p. 222). Therefore, education is classified as not significant. On the individual level, both Edu2 and Edu3 are judged to be fragile by all criteria and hence not significantly related to the probability of being risk-averse.

Concerning the sign of the relation, I find a strong tendency for a negative sign of the coefficient of both Edu2 and Edu3 that is even more pronounced for Edu2. This means, Heads who do not have a high school diploma (Edu1=1, reference category) have a higher probability of being risk-averse than Heads that have a high school diploma or an associate's degree (Edu2=1), and a higher probability of being risk-averse than Heads who have a bachelor's degree or higher (Edu3=1).

The slightly more pronounced negative effect of Edu2 compared to Edu3 to the probability of being risk-averse might be due to the definition of the three education levels, where Heads with high school diploma and those with an associate's degree are defined to belong to the same education level, namely Edu2. Since an associate's degree is often only an intermediate step towards higher education levels captured by Edu3, it might be that Heads with an associate's degree are more similar to Heads that already reached a higher education level (Edu3) than to those that only have a high school diploma (Edu2). If this was true, the effect of higher levels of education as defined by Edu3 is partly captured by Edu2.

Comparison to the literature and interpretation

Since the effect of education on risk-attitudes found in literature is ambiguous, my findings are partly in line with this literature as follows:[294] A negative effect of education on risk-aversion in the non-migration domain is found by Donkers, Melenberg, and van Soest (2001), Bonin, Constant, Tatsiramos, Zimmermann (2006), Jaeger, Bonin, Dohmen, Falk, Huffman, and Sunde (2007), Jaeger, Dohmen, Falk, Huffman, Sunde and Bonin (2008), and Umblijs (2012). The opposite effect is found by

[294] For a detailed discussion of findings on the relation for education and risk-attitudes in the previous empirical literature please refer to Part C, p.168.

Harrison, Lau, and Rutström (2007), and no significant effect at all by Hartog, Ferrer-i-Carbonell, and Jonker (2002).

Irrespective of findings in the literature, the strong tendency for a negative relation between education and the probability of being risk-averse in the migration context is what I expected to find since I expect people with higher education levels to be more capable of realizing and evaluating the risk involved in the migration decision. This might give them the feeling of controlling the risk, resulting in a higher willingness to take risks. Quite the opposite might hold for less educated people. They might see the risk but are not able to control it and, therefore, more reluctant to take it.

4.3.4.3 Detailed statistical reasoning

Results on joint significance of Edu2 and Edu3 are reported in Table 39, p. 222. Detailed statistical results from the extreme bounds analysis on the influence of education dummies Edu2 and Edu3 on the probability of being risk-averse are summarized separately for each dummy variable in Table 40, p. 223 and Table 41, p. 224, respectively.

(1)	(2)	(3)	(4)	(5)	(6)
Group of ways to estimate risk-attitude	Number of dependent variables analyzed	Missing values due to (quasi-) complete separation	Percentage of n where Edu2 and Edu3 are jointly significant at level p for 95% or more of the regressions run on that dependent variables		
			p=10%	p=5%	p=1%
	n	% of 4,536	% of n	% of n	% of n
Total	4,485	1%	38%	37%	36%
Ind	3,858	1%	37%	36%	35%
Fam	627	3%	42%	40%	39%
Ann	1,495	1%	38%	38%	36%
Wor	1,495	1%	37%	37%	35%
Lif	1,495	1%	37%	37%	35%
Var	2,268	0%	14%	14%	13%
LP2	2,217	2%	62%	61%	58%
Ed1	1,272	2%	39%	38%	36%
Ed2	645	0%	37%	36%	35%
Ed3	636	2%	35%	34%	33%
Ed4	636	2%	38%	37%	36%
Ed5	648	0%	37%	37%	36%
Ed6	648	0%	38%	37%	37%
Wei	2,229	2%	48%	47%	45%
Unw	2,256	1%	27%	27%	26%
Sep	2,256	1%	47%	45%	44%
Poo	2,229	2%	29%	28%	27%
One	1,512	0%	45%	45%	44%
Ad1	1,491	1%	33%	33%	31%
Ad2	1,482	2%	34%	33%	32%
L1	1,509	0%	32%	32%	31%
L2	1,506	0%	50%	50%	47%
Ma	1,470	3%	31%	29%	28%
2Kat3	1,512	0%	27%	26%	26%
2Kat1	1,500	1%	39%	39%	38%
4Kat2	1,473	3%	47%	46%	42%

Table 39: Percentages of dependent variables where Edu2 and Edu3 are jointly significant at level p for 95% or more of the regressions aggregated over all 4,536 dependent variables and subgroups of dependent variables.
The joint significance of Edu2 and Edu3 was tested using the Wald-test.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,485	1%	30%	0%	(-)	40%	7%	(-)	18%	(-)	42%	2%	(-)	15%	(-)
Ind	3,858	1%	30%	0%	(-)	40%	7%	(-)	18%	(-)	41%	2%	(-)	14%	(-)
Fam	627	3%	30%	0%	(-)	41%	6%	(-)	19%	(-)	47%	3%	(-)	18%	(-)
Ann	1,495	1%	30%	0%	(-)	41%	7%	(-)	18%	(-)	41%	2%	(-)	15%	(-)
Wor	1,495	1%	29%	0%	(-)	40%	7%	(-)	17%	(-)	42%	2%	(-)	15%	(-)
Lif	1,495	1%	29%	0%	(-)	41%	7%	(-)	18%	(-)	42%	2%	(-)	15%	(-)
Var	2,268	0%	13%	0%	(-)	20%	3%	(-)	15%	(-)	20%	1%	(-)	15%	(-)
LP2	2,217	2%	46%	0%	(-)	61%	8%	(-)	21%	(-)	64%	2%	(-)	15%	(-)
Ed1	1,272	2%	30%	0%	(-)	41%	5%	(-)	17%	(-)	44%	1%	(-)	15%	(-)
Ed2	645	0%	30%	0%	(-)	40%	9%	(-)	23%	(-)	40%	5%	(-)	19%	(-)
Ed3	636	2%	28%	0%	(-)	37%	5%	(-)	18%	(-)	38%	0%	(-)	16%	(-)
Ed4	636	2%	31%	0%	(-)	42%	5%	(-)	11%	(-)	42%	0%	(-)	9%	(-)
Ed5	648	0%	29%	0%	(-)	41%	7%	(-)	16%	(-)	42%	0%	(-)	11%	(-)
Ed6	648	0%	29%	0%	(-)	41%	11%	(-)	22%	(-)	41%	4%	(-)	21%	(-)
Wei	2,229	2%	39%	0%	(-)	50%	1%	(-)	9%	(-)	51%	0%	(-)	7%	(-)
Unw	2,256	1%	21%	0%	(-)	31%	16%	(-)	26%		32%	4%	(-)	23%	(-)
Sep	2,256	1%	37%	0%	(-)	48%	3%	(-)	14%	(-)	51%	2%	(-)	11%	(-)
Poo	2,229	2%	22%	0%	(-)	32%	12%	(-)	21%	(-)	32%	1%	(-)	19%	(-)
One	1,512	0%	35%	0%	(-)	47%	4%	(-)	19%	(-)	47%	3%	(-)	16%	(-)
Ad1	1,491	1%	26%	0%	(-)	36%	10%	(-)	18%	(-)	38%	2%	(-)	15%	(-)
Ad2	1,482	2%	27%	0%	(-)	38%	7%	(-)	16%	(-)	39%	0%	(-)	14%	(-)
L1	1,509	0%	30%	0%	(-)	37%	0%	(-)	1%	(-)	39%	0%	(-)	1%	(-)
L2	1,506	0%	41%	0%	(-)	50%	0%	(-)	8%	(-)	52%	0%	(-)	8%	(-)
Ma	1,470	3%	17%	0%	(-)	34%	25%	(-)	44%		33%	7%	(-)	36%	
2Kat3	1,512	0%	23%	0%	(-)	28%	6%	(-)	19%	(-)	28%	2%	(-)	17%	(-)
2Kat1	1,500	1%	32%	0%	(-)	41%	7%	(-)	18%	(-)	40%	1%	(-)	13%	(-)
4Kat2	1,473	3%	34%	0%	(-)	52%	7%	(-)	16%	(-)	57%	2%	(-)	14%	(-)

Table 40: Results of the extreme bounds analysis for the independent variable Edu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a tendency for a positive coefficient, (-) denotes a tendency for a positive coefficient.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,485	1%	20%	0%	(-)	48%	14%	(-)	41%		44%	13%	(-)	43%	
Ind	3,858	1%	21%	0%	(-)	47%	13%	(-)	42%		43%	12%	(-)	43%	
Fam	627	3%	18%	0%	(-)	54%	23%	(-)	39%		52%	17%	(-)	41%	
Ann	1,495	1%	21%	0%	(-)	48%	14%	(-)	41%		45%	13%	(-)	43%	
Wor	1,495	1%	20%	0%	(-)	48%	14%	(-)	41%		44%	13%	(-)	43%	
Lif	1,495	1%	20%	0%	(-)	48%	15%	(-)	41%		44%	13%	(-)	43%	
Var	2,268	0%	10%	0%	(-)	29%	27%		58%		19%	26%		62%	
LP2	2,217	2%	30%	0%	(-)	68%	9%	(-)	25%	(-)	70%	10%	(-)	23%	(-)
Ed1	1,272	2%	18%	0%	(-)	51%	18%	(-)	40%		48%	16%	(-)	42%	
Ed2	645	0%	18%	0%	(-)	47%	16%	(-)	46%		46%	17%	(-)	47%	
Ed3	636	2%	15%	0%	(-)	47%	16%	(-)	48%		41%	16%	(-)	48%	
Ed4	636	2%	17%	0%	(-)	46%	8%	(-)	38%		42%	10%	(-)	43%	
Ed5	648	0%	28%	0%	(-)	47%	9%	(-)	36%		39%	3%	(-)	37%	
Ed6	648	0%	28%	0%	(-)	49%	16%	(-)	41%		44%	12%	(-)	41%	
Wei	2,229	2%	27%	0%	(-)	55%	6%	(-)	23%	(-)	51%	6%	(-)	25%	(-)
Unw	2,256	1%	13%	0%	(-)	41%	26%		60%		38%	23%	(-)	61%	
Sep	2,256	1%	24%	0%	(-)	54%	7%	(-)	38%		51%	5%	(-)	40%	
Poo	2,229	2%	16%	0%	(-)	42%	24%	(-)	45%		37%	25%	(-)	46%	
One	1,512	0%	36%	0%	(-)	52%	7%	(-)	32%		46%	3%	(-)	29%	
Ad1	1,491	1%	12%	0%	(-)	46%	18%	(-)	44%		41%	17%	(-)	50%	
Ad2	1,482	2%	13%	0%	(-)	47%	20%	(-)	48%		45%	20%	(-)	50%	
L1	1,509	0%	17%	0%	(-)	47%	21%	(-)	49%		39%	19%	(-)	49%	
L2	1,506	0%	30%	0%	(-)	54%	4%	(-)	36%		56%	6%	(-)	36%	
Ma	1,470	3%	15%	0%	(-)	44%	21%	(-)	39%		37%	17%	(-)	43%	
2Kat3	1,512	0%	19%	0%	(-)	36%	21%	(-)	53%		34%	22%	(-)	52%	
2Kat1	1,500	1%	23%	0%	(-)	45%	11%	(-)	43%		43%	12%	(-)	45%	
4Kat2	1,473	3%	20%	0%	(-)	64%	13%	(-)	28%		56%	8%	(-)	31%	

Table 41: Results of the extreme bounds analysis for the independent variable Edu3 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a tendency for a positive coefficient, (-) denotes a tendency for a positive coefficient.

Sample size and (quasi-) complete separation

Since (quasi-) complete separation only occurs for 1% of the 4,536 dependent variables (Column (3), Table 39, p. 222) the statistical reasoning on robustness, sign, and significance of both education dummies Edu2 and Edu3 is based on a sufficiently high sample of 4,485 dependent variables (Column (2), Table 39, p. 222).

Joint significance of Edu2 and Edu3

Education is classified as not significant since Edu2 and Edu3 are jointly significant at 1% significance level for only 36% of the 4,485 dependent variables analyzed (Column (4), Table 39, p. 222). If the required significance level is further reduced to p=10%, a significant influence of education on the probability of being risk-averse (Column (6), Table 39, p. 222) is only found for 38% of the 4,485 dependent variables is found.

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, Edu2 and Edu3 are individually classified as fragile: Edu2 exhibits a robust negative influence on 30% of the 4,485 dependent variables (Column (4), Table 40, p. 223), and Edu3 exhibits a robust negative influence on 20% of the 4,485 dependent variables (Column (4), Table 41, p. 224). Hence, Edu2 and Edu3 are individually not significantly related to the probability of being risk-averse in the migration context.

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, Edu2 and Edu3 are also classified as fragile. Of the 4,485 dependent variables percentages of robust relations range from 40% by Sala-I-Martin's weighted criterion on Edu2 (Column (7), Table 40, p. 223) to 48% by Sala-I-Martin's weighted criterion on Edu3 (Column (7), Table 41, p. 224).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, results indicate a strong tendency for a negative sign for both Edu2 and Edu3 and their relation to the probability of being risk-averse, where the tendency for a negative sign is even more pronounced for Edu2. Put differently, Heads who do not have a high school diploma (Edu1=1) have a higher probability of being risk-averse than Heads that have a high school diploma or an associate's degree (Edu2=1), and a higher probability of being risk-averse than Heads who have a bachelor's degree or higher (Edu3=1).

4.3.4.4 Sensitivity for subgroups of dependent variables

This section checks whether the statistical statements derived for education variables Edu2 and Edu3 systematically differ for certain subgroups of dependent variables, where the sensitivity analysis is run first, on joint significance of Edu2 and Edu3, and second, separately for Edu2 and Edu3.

4.3.4.4.1 Overview

Joint significance of Edu2 and Edu3

The non-significance of education in explaining the probability of being risk-averse over the 4,485 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy. First, comparing results of subgroups referring to different risk-measures, I find that the percentage of dependent variables for which Edu2 and Edu3 are jointly significant at a 1% significance level in the subgroup referring to semi-variance (LP2) is more than four times the percentage in the subgroup referring to variance (Var). Second, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of dependent variables for which Edu2 and Edu3 are jointly significant at a 1% significance level increase visibly.

Robustness

The fragility of education in explaining the probability of being risk-averse over the 4,485 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing results of subgroups referring to different risk-measures, I find that the subgroup referring to risk-measure semi-variance (LP2) exhibits more than three times the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria than the subgroup referring to variance (Var). Second, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted increase visibly.

Tendency for a sign

The finding of a strong tendency for a negative sign of Edu2 for the 4,485 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In contrast, the strong tendency for a negative sign of Edu3 and its relation to the probability of being risk-averse that was found for the 4,485 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the education definition (denoted by Ed1 to Ed6), (iv) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (v) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), and not sensitive to (vi) variations in the measurement of predictive errors (denoted by L1, L2, and Ma).

However, the finding of a strong tendency for a negative sign of Edu3 is sensitive to (i) variations of the risk-measure (denoted by LP2 and Var), (ii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), and sensitive to (iii) variations of the transformation rule (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.4.4.2 Detailed statistical reasoning and interpretation

I first investigate the joint significance of Edu2 and Edu3 for its sensitivity to variations of the way risk-attitudes are estimated. I then turn to the sensitivity analysis separately for Edu2 and Edu3.

Two facts are identical for both Edu2 and Edu3 over all subgroups and will therefore not be mentioned again for every subgroup. That is, first, the percentages of (quasi-) complete separation are equally low ranging from 0% to 3% of the dependent variables in all subgroups, and are therefore of no concern for any subgroup. Second, both Edu2 and Edu3 are individually not significantly related to the probability of being risk-averse in any subgroup. Consequently, only statements on robustness and sign are investigated for their sensitivity separately by Edu2 and Edu3 as follows:

<u>Detailed comparison of subgroups of dependent variables relating to different decision problems</u>

- The relation of education to the probability of being risk-averse is not sensitive to whether the migration decision is estimated based on **Head's individual income or family income** (denoted by Ind and Fam). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income only education definition one was applied. Therefore, the number of dependent variables in the subgroup referring to family income (Fam) amounts to one sixth of the dependent variables based in Head's individual income (Ind).
    - **Joint significance of Edu2 and Edu3:** In both subgroups Edu2 and Edu3 are not jointly significant at a 10% level.

- **Sensitivity of Edu2:** For both subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, both subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Edu3:** For both subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, both subgroups show a strong tendency for a negative sign of the coefficient of Edu3.

I interpret these findings as follows: Irrespective of whether Head's individual income or family income is considered, compared to Heads with higher levels of education, Heads with lower levels of education exhibit a higher probability of being risk-averse towards variations of their own individual income as well as towards variations of family income.

- The relation of education to the probability of being risk-averse is not sensitive to **different planning periods** investigated in my study, i.e., one year, time until reaching full retirement age, and time until reaching life expectancy (denoted by Ann, Wor, and Lif).

    - **Joint significance of Edu2 and Edu3:** In all subgroups Edu2 and Edu3 are not jointly significant at a 10% level.

    - **Sensitivity of Edu2:** For all subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

    - **Sensitivity of Edu3:** For all subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu3.

I interpret these findings as being due to the way income parameters for longer planning periods have been estimated (for a detailed discussion of this argument see the corresponding interpretation for Male, p. 204).

- The relation of education to the probability of being risk-averse is sensitive to **different risk-measures** variance and semi-variance (denoted by Var and LP2).

    - **Joint significance of Edu2 and Edu3:** Although in both subgroups Edu2 and Edu3 are not jointly significant at a 10% level, it is noteworthy that in the subgroup referring to semi-variance (LP2) the percentage of dependent variables where Edu2 and Edu3 are jointly significant is more than four times the percentage in the subgroup referring to variance (Var) for significance levels p=1%, p=5%, and p=10% (Columns (4) to (6), Table 39, p. 222).

- **Sensitivity of Edu2:** First, although Edu2 is fragile by Leamer's criterion in both subgroups, it is noteworthy that in the subgroup referring to semi-variance (LP2) the percentage of robust relations is about three times the percentage in the subgroup referring to variance (Var). Second, applying Sala-I-Martin's criterion, Edu2 is again classified as fragile where the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about two times (three times) the percentage in the subgroup referring to variance (Var). Third, both subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Ed3:** First, although Edu3 is fragile by Leamer's criterion in both subgroups, it is noteworthy that in the subgroup referring to semi-variance (LP2) the percentage of robust relations is about three times the percentage in the subgroup referring to variance (Var). Second, applying Sala-I-Martin's criterion, Edu3 is again classified as fragile where the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about two times (three times) the percentage in the subgroup referring to variance (Var). Third, for Edu3 I find different tendencies for a sign in both subgroups, i.e., the subgroup referring to variance (Var) shows no tendency for a sign, while the subgroup referring to semi-variance (LP2) shows a strong tendency for a negative sign of the coefficient of Edu3.

I interpret these findings as follows: First, the stronger relation (as indicated by (i) a higher percentages of dependent variables where Edu2 and Edu3 are jointly significant and (ii) a higher percentages of robust relations by all criterion for the subgroup referring to semi-variance compared to the subgroup referring to variance) of education to the probability of being risk-averse towards the downside-risk (LP2) indicates that Heads of different education levels are more different in their risk-attitude towards the downside-risk (LP2) than they are towards the risk of any variation of income (Var). A possible explanation might be that Heads with higher levels of education compared to Heads with lower level of education are more confident in their skills and therefore less afraid of earning less than expected, while they do not differ so much in their attitude towards any variation of income. Second, the fact that a tendency for sign cannot be observed for the highest education level (Edu3) in the subgroup referring to variance (Var) might be due a combination of a generally weaker relation between education and probability of being risk-averse for (i) risk-measure variance (Var) compared to semi-variance (LP2), and for (ii) Edu3 compared to Edu2.

<u>Detailed comparison of subgroups of dependent variables relating to different ways to gain data input</u>

- The relation of education to the probability of being risk-averse is not sensitive to **different education definitions** (denoted by Ed1 to Ed6). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income only education definition one was applied. Therefore, the number of dependent variables to be analyzed amounts to 648 for education definitions two to six, while education definition one additionally includes all migration decisions based on family income.

    - **Joint significance of Edu2 and Edu3:** In all subgroups Edu2 and Edu3 are not jointly significant at a 10% level.

    - **Sensitivity of Edu2:** For all subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

    - **Sensitivity of Edu3:** For all subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu3.

  I interpret these findings as follows: Since the six education definitions are due to competing answers on (1) whether education variables from years before or after the move should be preferred, and (2) how to treat opposing education information in the family and individual data file, I conclude that the relation of education to the probability of being risk-averse is not sensitive to competing answers on both questions.

- The relation of education to the probability of being risk-averse is sensitive to whether the **income parameters are estimated based on weighted or unweighted samples** (denoted by Wei and Unw).

    - **Joint significance of Edu2 and Edu3:** Although in both subgroups Edu2 and Edu3 are not jointly significant at a 10% level, it is noteworthy that in the subgroup referring to weighted samples (Wei) the percentage of dependent variables where Edu2 and Edu3 are jointly significant is about twice the percentage in the subgroup referring to unweighted samples (Unw) for significance levels p=1%, p=5%, and p=10% (Columns (4) to (6), Table 39, p. 222).

- **Sensitivity of Edu2:** First, although Edu2 is fragile by Leamer's criterion in both subgroups, it is noteworthy that in the subgroup referring to weighted samples (Wei) the percentage of robust relations is about 1.9 times the percentage in the subgroup referring to unweighted samples (Unw). Second, by Sala-I-Martin's criterion Edu2 is again fragile in both subgroups, but in the subgroup referring to weighted samples (Wei) the percentage of robust relations is about 1.6 times the percentage in the subgroup referring to unweighted samples (Unw). Third, both subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Ed3:** First, although Edu3 is fragile by Leamer's criterion in both subgroups, it is noteworthy that in the subgroup referring to weighted samples (Wei) the percentage of robust relations is about two times the percentage in the subgroup referring to unweighted samples (Unw). Second, by Sala-I-Martin's criterion Edu3 is again fragile in both subgroups, but in the subgroup referring to weighted samples (Wei) the percentage of robust relations is about 1.3 times the percentage in the subgroup referring to unweighted samples (Unw). Third, the subgroup referring to weighted samples (Wei) shows a strong tendency for a negative sign, while the subgroup referring to unweighted samples (Unw) shows no tendency for a sign.

I interpret these findings as follows: Obviously, it makes a difference for the influence of education on the probability of being risk-averse whether samples are weighted or not. Since both approaches are afflicted with problems, it is not clear whether it is (i) the unrepresentative sample of the subgroup referring to unweighted samples (Wei) or (ii) the weighting of the sample (Wei) that causes biased estimates of income parameters.

- The relation of education to the probability of being risk-averse is not sensitive to the **type of clustering** of people from which income parameters are estimated, i.e., separate clustering for each year versus pooled clustering with same clusters in all years (denoted by Sep and Poo).

    - **Joint significance of Edu2 and Edu3:** Although in both subgroups Edu2 and Edu3 are not jointly significant at a 10% level, it is noteworthy that in the subgroup referring to separate clustering (Sep) the percentage of dependent variables where Edu2 and Edu3 are jointly significant is about 1.6 times the percentage in the subgroup referring to pooled clustering (Poo) for significance levels p=1%, p=5%, and p=10% (Columns (4) to (6), Table 39, p. 222).

- **Sensitivity of Edu2:** For both subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion, where the percentages of robust relations are slightly higher by all criteria for the subgroup referring to separate clustering (Sep) compared to pooled clustering (Poo). Third, both subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Edu3:** For both subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion, where the percentages of robust relations are slightly higher by all criteria for the subgroup referring to separate clustering (Sep) compared to pooled clustering (Poo). Third, both subgroups show a strong tendency for a negative sign of the coefficient of Edu3.

I interpret these findings as follows: First, the non-sensitivity of the relation of education to the probability of being risk-averse is not surprising since different types of clustering result in the same education-clusters. Second, the considerably different age-clusters resulting from the two types of clustering do not seem to make a difference for the effect of education on the probability of being risk-averse. For example, depending on the cluster algorithm used it, could be that income parameters for a 30 years old men with a bachelor's degree are either (i) estimated from income data of all men with a bachelor's degree being 23 to 30 years old, or (ii) estimated from income data of all men with a bachelor's degree being 28 to 34 years old. It is therefore surprising that such great differences in the definitions of age-clusters do not alter the effect of education on the probability of being risk-averse in the migration context. Third, the slightly more pronounced results in the subgroup referring to separate clustering (Sep) might indicate that separate clustering results in more precise estimations of the income parameters on an annual basis (see Part B, Section 3.4.4.2).

- The relation of education to the probability of being risk-averse is not sensitive to different **time periods from which income parameters are estimated**, namely annual income data of one signle year, annual income date from three years around the moving date adjusted for inflation, and annual income data from three years around the moving date not adjusted for inflation (denoted by One, Ad1, and Ad2).

- **Joint significance of Edu2 and Edu3:** In all subgroups Edu2 and Edu3 are not jointly significant at a 10% level.

- **Sensitivity of Edu2:** For all subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Edu3:** For all subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu3.

I interpret these findings as follows: It seems as if the results of education and its relation to the probability of being risk-averse are not sensitive to whether estimations are based on one year data compared to three years of data, and also not sensitive to whether the three year's data are adjusted for inflation or not.

<u>Detailed comparison of subgroups of dependent variables relating to different estimation procedures of risk-attitudes</u>

The relation of education to the probability of being risk-averse is not sensitive to **different measurements of predictive errors**, namely L_p-norms one-, two- and infinity-norms where $p = \{1,2,\infty\}$ (denoted by L1, L2, and Ma).

- **Joint significance of Edu2 and Edu3:** In all subgroups Edu2 and Edu3 are not jointly significant at a 10% level.

- **Sensitivity of Edu2:** For all subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Edu3:** For all subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu3.

I interpret these findings as follows: It seems as if the maximum predictive error is decisive since results for subgroups that additionally account for smaller predictive errors (L1 and L2) exhibit qualitative the same results.

<u>Detailed comparison of subgroups of dependent variables relating to different transformation rules to code the binary dependent variables</u>

The relation of education to the probability of being risk-averse is sensitive to **different transformation rules** (denoted by 2Kat1, 2Kat3, and 4Kat2).

- **Joint significance of Edu2 and Edu3:** Although in all subgroups Edu2 and Edu3 are not jointly significant at a 10% level, it is noteworthy that the percentage of dependent variables where Edu2 and Edu3 are jointly significant at any significance level surveyed rises when Heads with less pronounced degrees of risk-attitude are deleted from the sample step by step (i.e., 2Kat3, 2Kat1, and 4Kat2).

- **Sensitivity of Edu2:** For all subgroups I find that Edu2 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Concerning robustness, it is again noteworthy that the percentage of robust relations by any criterion rises when migrants with lower degrees of risk-attitude are deleted from the sample step by step. Third, all subgroups show a strong tendency for a negative sign of the coefficient of Edu2.

- **Sensitivity of Edu3:** For all subgroups I find that Edu3 is classified as fragile by first, Leamer's criterion, and second, Sala-I-Martin's criterion. Concerning robustness, it is again noteworthy that the percentage of robust relations by any criterion rises when migrants with lower degrees of risk-attitude are deleted from the sample step by step. Third, concerning the tendency for a sign, I find no tendency for the subgroup that includes all migrants (2Kat3), but a strong tendency for a negative sign for subgroups that delete migrants with lower degrees of risk-attitude from the sample (2Kat12 and 4Kat2).

I interpret these findings as follows: The idea of the extreme transformation rule (4Kat2) was that once migrants with less pronounced degrees of risk-attitude are deleted from the sample, the true relation between an independent variable and the probability of being risk-averse would be more pronounced. This is exactly what can be seen for Edu2 and Edu3.

4.3.5 AgeSquared squared and risk-attitudes

4.3.5.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The relation of AgeSquared to the probability of being risk-averse can be statistically investigated for all 4,536 dependent variables since (quasi-) complete separation does not occur. Based on this sample, AgeSquared does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Therefore, a statement on significance cannot be made. Yet, I derive a strong tendency for a positive coefficient of AgeSquared. A positive quadratic age-effect means that the probability of being risk-averse increases at an increasing rate as Age increases when people turn older if there is no additional linear age effect.

Although AgeSquared is fragile, the strong tendency for a positive sign might explain why an isolated linear age-effect as captured by the independent variable Age was not found (for details on this argument see Part C, Section 4.3.3). Still, a total age-effect that combines linear and quadratic age-effects is not estimated in my study for the following reason: The extreme bounds analysis used in my study aggregates regression results over several regressions. For each regression the total age-effect most certainly has another extremum and maybe even a different shape (for example, U-shaped versus inverse U-shaped). Therefore, aggregating results of numerous regressions to gain results for one dependent variable - as done in my study - would most certainly make a sensible interpretation in economic terms impossible. For example, if in the first regression the maximum turning point is reached when AgeSquared equals 5, and in a second regression when AgeSquared equals 65, the interpretation of both total age-effects contradicts each other. The total age-effect of the first regression indicates that the probability of being risk-averse is constantly decreasing when Heads turn older. Recall that Heads younger than 5 years do not exist in my sample. In contrast, the total age-effect of the second regression indicates that the probability of being risk-averse is constantly increasing when Heads turn older until they reach the age of 65. Recall, that the majority of Heads in my sample is younger than 65. Obviously, results of both regressions contradict each other in their economic interpretation. Averaging the total age-effect over all regressions would consequently be misleading since it is not clear which regression actually captures the true total age-effect.

Comparison to the literature and interpretation

My findings regarding AgeSquared are partly in line with the literature as follows: [295] On the one hand, the strong tendency for a positive sign of AgeSquared is in line with findings of Hallahan, Faff, and McKenzie (2004) in the domain of finance, Bonin, Constant, Tatsiramos, and Zimmermann (2006) for the domains of finance, car driving, and career, and Säve-Söderbergh (2012) in the domain of finance. On the other hand, the fragility of AgeSquared in explaining the probability of being risk-averse contradicts previous findings in the non-migration literature where the quadratic AgeSquared term was found to have a significant influence.

The difference of my findings compared to those of the literature is not surprising to me for two reasons. First, risk-attitudes are domain-specific and the domain of migration has not been surveyed in the literature yet. Second, empirical studies that include a quadratic age-effect also include a linear age-effect at the same time. Results of these studies are therefore not comparable to an isolated quadratic age-effect.

4.3.5.2 Detailed statistical reasoning

All statistical results on the influence of AgeSquared on the probability of being risk-averse are summarized in Table 42, p. 237.

[295] For a detailed discussion of findings on the relation of age and age squared on risk-attitudes in the previous empirical literature please refer to Part C, p. 169.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	**number of dependent variables**	**missing values**	**Leamer**			**Sala-I-Martin weighted**					**Sala-I-Martin unweighted**				
			significant	**positive**		**CDF(0) ≥95%**	**positive**		**positive**		**CDF(0) ≥95%**	**positive**		**positive**	
	n	**% of 4,536**	**% of n**	**% of significant**		**% of n**	**% of robust**		**% of n**		**% of n**	**% of robust**		**% of n**	
Total	4,536	0%	0.3%	100%	(+)	21%	89%	(+)	77%	(+)	14%	88%	(+)	79%	(+)
Ind	3,888	0%	0.4%	100%	(+)	22%	89%	(+)	79%	(+)	15%	87%	(+)	81%	(+)
Fam	648	0%	0.0%	n.a.		14%	88%	(+)	70%		11%	96%	(+)	71%	
Ann	1,512	0%	0.3%	100%	(+)	20%	89%	(+)	76%	(+)	14%	88%	(+)	79%	(+)
Wor	1,512	0%	0.3%	100%	(+)	21%	89%	(+)	78%	(+)	15%	88%	(+)	79%	(+)
Lif	1,512	0%	0.3%	100%	(+)	21%	88%	(+)	78%	(+)	15%	87%	(+)	80%	(+)
Var	2,268	0%	0.0%	n.a.		9%	56%		67%		7%	47%		68%	
LP2	2,268	0%	0.7%	100%	(+)	33%	97%	(+)	87%	(+)	22%	100%	(+)	91%	(+)
Ed1	1,296	0%	0.2%	100%	(+)	19%	90%	(+)	76%	(+)	14%	92%	(+)	79%	(+)
Ed2	648	0%	0.5%	100%	(+)	23%	91%	(+)	80%	(+)	17%	91%	(+)	85%	(+)
Ed3	648	0%	0.5%	100%	(+)	18%	85%	(+)	70%		13%	80%	(+)	70%	
Ed4	648	0%	0.5%	100%	(+)	19%	84%	(+)	73%		13%	83%	(+)	71%	
Ed5	648	0%	0.5%	100%	(+)	23%	90%	(+)	84%	(+)	16%	84%	(+)	85%	(+)
Ed6	648	0%	0.0%	n.a.		23%	89%	(+)	82%	(+)	14%	89%	(+)	85%	(+)
Wei	2,268	0%	0.7%	100%	(+)	22%	98%	(+)	82%	(+)	17%	97%	(+)	82%	(+)
Unw	2,268	0%	0.0%	n.a.		19%	78%	(+)	72%		11%	73%		77%	(+)
Sep	2,268	0%	0.0%	n.a.		16%	93%	(+)	81%	(+)	13%	93%	(+)	80%	(+)
Poo	2,268	0%	0.7%	100%	(+)	25%	86%	(+)	74%		16%	84%	(+)	78%	(+)
One	1,512	0%	1.0%	100%	(+)	29%	91%	(+)	88%	(+)	26%	91%	(+)	90%	(+)
Ad1	1,512	0%	0.0%	n.a.		15%	86%	(+)	70%		9%	79%	(+)	71%	
Ad2	1,512	0%	0.0%	n.a.		17%	86%	(+)	75%	(+)	8%	87%	(+)	77%	(+)
L1	1,512	0%	0.0%	n.a.		11%	100%	(+)	84%	(+)	8%	99%	(+)	82%	(+)
L2	1,512	0%	0.0%	n.a.		13%	97%	(+)	73%		10%	100%	(+)	84%	(+)
Ma	1,512	0%	1.0%	100%	(+)	38%	82%	(+)	76%	(+)	26%	79%	(+)	72%	
2Kat3	1,512	0%	1.0%	100%	(+)	23%	75%	(+)	74%		16%	74%		81%	(+)
2Kat1	1,512	0%	0.0%	n.a.		15%	96%	(+)	78%	(+)	12%	96%	(+)	81%	(+)
4Kat2	1,512	0%	0.0%	n.a.		25%	97%	(+)	80%	(+)	15%	95%	(+)	76%	(+)

Table 42: Results of the extreme bounds analysis for the independent variable AgeSquared aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Sample size and (quasi-) complete separation

(Quasi-) Complete separation does not occur. Hence, the statistical reasoning on robustness, sign, and significance of AgeSquared is based on all 4,536 dependent variables (Column (2), Table 42, p. 237).

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, AgeSquared is clearly classified as fragile, i.e., AgeSquared does only show a robust significant influence on 0.3% of the 4,536 dependent variables (Column (4), Table 42, p. 237). Consequently, AgeSquared is classified as not significantly related to probability of being risk-averse in the migration context.

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, AgeSquared is also clearly classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 21% (14%) of the 4,536 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 42, p. 237).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, I observe a strong tendency for a positive sign of the coefficient of AgeSquared.

4.3.5.3 Sensitivity for subgroups of dependent variables

4.3.5.3.1 Overview

Robustness

The fragility and non-significance of AgeSquared in explaining the probability of being risk-averse over all 4,536 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, one finding is noteworthy: Comparing results of subgroups referring to different risk-measures, I find that the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria in the subgroup referring to semi-variances (LP2) are more than three times the percentages of robust relations in the subgroup referring to variance (Var).

Tendency for a sign

The finding of a strong tendency for a positive sign of AgeSquared for all 4,536 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the education definition (denoted by Ed1 to Ed6), (iv) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (v) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (vi) whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2), (vii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (viii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

However, the finding of a strong tendency for a positive sign of AgeSquared and its relation to the probability of being risk-averse for all 4,536 dependent variables statistically investigated is sensitive to variations of the risk-measure (denoted by Var and LP2).

4.3.5.3.2 Detailed statistical reasoning and interpretation

Since the effect of AgeSquared on the probability of being risk-averse is only sensitive to different risk-measures, I only discuss this very sensitivity in detail to avoid unnecessary repetitions. Numerical details on all subgroups are given in Table 42, p. 237.

Detailed comparison of subgroups of dependent variables relating to different decision problems

The relation of AgeSquared to the probability of being risk-averse is sensitive to **different risk-measures** variance and semi-variance (denoted by Var and LP2). First, (quasi-) complete separation does not occur in any subgroup. Second, applying Leamer's criterion of robustness, AgeSquared is classified as fragile in both subgroups. Third, therefore AgeSquared is also classified as not significantly related to probability of being risk-averse in both subgroups. Fourth, applying Sala-I-Martin' criterion of robustness, AgeSquared is classified as fragile in both subgroups, but in the subgroup referring to semi-variance (LP2) the percentages of robust relations is about two times the percentages in the subgroup referring to variance (Var). Fifth, concerning the tendency for a sign of the coefficient of AgeSquared, I observe no tendency for a sign in the subgroup referring to variance (Var) and a strong tendency for a negative sign in the subgroup referring to semi-variance (LP2).

I interpret the finding of (i) a lower percentage of robust relations in the subgroup referring to variance (Var) together with (ii) the non-existing tendency for a sign in the subgroup referring to variance (Var) as follows: First, if the total age-effect consists of a linear and a quadratic age-effect, a possible interpretation might be that the probability of being risk-averse towards the downside-risk (LP2) increases at an increasing rate when people turn older, while the effect of an additional year of age on the probability of being risk-averse towards variance of income (Var) seems to be not quadratic. Second, if the insights from interpretation of the linear age-effect (Age) and the quadratic age-effect (AgeSquared) are combined, the total age-effect for risk-measure variance (Var) must be a polynomial of higher degree.

4.3.6 Family size and risk-attitudes

To measure the effect of family size on Head's risk-attitude in the domain of migration, two proxies are available, namely the number of family members before the move and the number of family members after the move. Since it is not clear which of the variables is the right proxy for family size, results on both variables are discussed in the following sections.

4.3.6.1 Number of family members in the wave before the move

4.3.6.1.1 Overview of results irrespective of the way risk-attitudes are estimated

<u>Significance, robustness, and tendency for a sign</u>

The influence of the number of family members before the move on the probability of being risk-averse can be statistically investigated for all 4,536 dependent variables since (quasi-) complete separation does not occur. Based on this sample, the number of family members before the move does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Consequently, the number of family members in the wave before the move is also not significantly related to the probability of being risk-averse in the migration context. Yet, I derive a strong tendency for a positive sign of the number of family members before the move, which means the probability of being risk-averse is higher for migrants that lived in a greater family before the move compared to migrants that lived in a smaller family before the move. This might be due to the increasing financial responsibility Heads have to bear when the number of family members before the move increases.

Comparison to the literature

My findings are partly in line with the literature on family size and risk-attitudes as follows:[296] On the one hand, no significant effect of family size on risk-attitudes in the non-migration domain is found by Harrison, Lau, and Rutström (2007) and Säve-Söderbergh (2012), while most other authors find a significant effect of family size on risk-attitudes in the non-migration domain. On the other hand, the strong tendency for a positive sign is in line with previous findings in developed countries where Heads of bigger families showed significantly higher risk-aversion in the non-migration domain compared to Heads of smaller families.

The difference of my findings compared to those of the literature is not surprising to me for two reasons. First, risk-attitudes are domain-specific and the domain of migration has not been surveyed yet. Second, the number of family members in the wave before the move might not fully capture the situation at the time of the move since the number of family members in the wave before the move does not necessarily relate to the number of family members that actually move together with Head.

4.3.6.1.2 Detailed statistical reasoning

All statistical results on the influence of the number of family members in the wave before the move on the probability of being risk-averse are summarized in Table 43, p. 242.

[296] For a detailed discussion of findings on the relation for family size and risk-attitudes in the previous empirical literature please refer to Part C, p. 170.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,536	0%	1.1%	100%	(+)	27%	99%	(+)	81%	(+)	19%	99%	(+)	82%	(+)
Ind	3,888	0%	1.3%	100%	(+)	27%	98%	(+)	81%	(+)	20%	99%	(+)	82%	(+)
Fam	648	0%	0.0%	n.a.		23%	100%	(+)	80%	(+)	13%	100%	(+)	85%	(+)
Ann	1,512	0%	1.1%	100%	(+)	27%	99%	(+)	82%	(+)	19%	99%	(+)	83%	(+)
Wor	1,512	0%	1.1%	100%	(+)	26%	98%	(+)	81%	(+)	19%	99%	(+)	83%	(+)
Lif	1,512	0%	1.1%	100%	(+)	26%	98%	(+)	81%	(+)	19%	99%	(+)	82%	(+)
Var	2,268	0%	0.3%	100%	(+)	12%	100%	(+)	70%		8%	100%	(+)	71%	
LP2	2,268	0%	2.0%	100%	(+)	42%	98%	(+)	93%	(+)	30%	99%	(+)	94%	(+)
Ed1	1,296	0%	0.9%	100%	(+)	27%	99%	(+)	83%	(+)	17%	99%	(+)	85%	(+)
Ed2	648	0%	0.9%	100%	(+)	29%	95%	(+)	79%	(+)	19%	98%	(+)	81%	(+)
Ed3	648	0%	0.5%	100%	(+)	18%	100%	(+)	77%	(+)	15%	100%	(+)	76%	(+)
Ed4	648	0%	0.5%	100%	(+)	18%	100%	(+)	76%	(+)	15%	100%	(+)	76%	(+)
Ed5	648	0%	2.3%	100%	(+)	36%	97%	(+)	88%	(+)	27%	100%	(+)	90%	(+)
Ed6	648	0%	1.9%	100%	(+)	31%	100%	(+)	81%	(+)	24%	100%	(+)	84%	(+)
Wei	2,268	0%	0.0%	n.a.		12%	93%	(+)	66%		8%	97%	(+)	68%	
Unw	2,268	0%	2.2%	100%	(+)	42%	100%	(+)	96%	(+)	30%	100%	(+)	97%	(+)
Sep	2,268	0%	0.9%	100%	(+)	22%	100%	(+)	85%	(+)	17%	100%	(+)	85%	(+)
Poo	2,268	0%	1.3%	100%	(+)	31%	97%	(+)	78%	(+)	21%	99%	(+)	80%	(+)
One	1,512	0%	1.4%	100%	(+)	27%	100%	(+)	86%	(+)	24%	100%	(+)	89%	(+)
Ad1	1,512	0%	1.2%	100%	(+)	26%	98%	(+)	80%	(+)	16%	100%	(+)	81%	(+)
Ad2	1,512	0%	0.8%	100%	(+)	28%	97%	(+)	78%	(+)	18%	98%	(+)	78%	(+)
L1	1,512	0%	2.2%	100%	(+)	32%	100%	(+)	84%	(+)	29%	100%	(+)	86%	(+)
L2	1,512	0%	0.8%	100%	(+)	27%	100%	(+)	82%	(+)	14%	100%	(+)	84%	(+)
Ma	1,512	0%	0.4%	100%	(+)	20%	94%	(+)	78%	(+)	15%	97%	(+)	78%	(+)
2Kat3	1,512	0%	3.4%	100%	(+)	36%	100%	(+)	87%	(+)	29%	100%	(+)	88%	(+)
2Kat1	1,512	0%	0.0%	n.a.		27%	99%	(+)	83%	(+)	21%	98%	(+)	86%	(+)
4Kat2	1,512	0%	0.0%	n.a.		17%	95%	(+)	73%		7%	100%	(+)	74%	

Table 43: Results of the extreme bounds analysis of the independent variable “number of family members in the wave before the move” aggregated over all 4,536 dependent variables and subgroups of dependent variables.
Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

<u>Sample size and (quasi-) complete separation</u>

(Quasi-) Complete separation does not occur. Hence, the statistical reasoning on robustness, sign, and significance of the number of family members in the wave before the move is based on all 4,536 dependent variables (Column (2), Table 43, p. 242).

<u>Robustness and significance based on Leamer's criterion</u>

Applying Leamer's criterion of robustness, the number of family members in the wave before the move is clearly classified as fragile, i.e., the number of family members in the wave before the move does only show a robust significant influence on 1.1% of the 4,536 dependent variables (Column (4), Table 43, p. 242). Consequently, the number of family members in the wave before the move is classified as not significantly related to probability of being risk-averse in the migration context.

<u>Robustness based on Sala-I-Martin's criterion</u>

Applying Sala-I-Martin's criterion of robustness, the number of family members in the wave before the move is classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 27% (19%) of the 4,536 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 43, p. 242).

<u>Tendency for a sign based on Leamer's and Sala-I-Martin's criteria</u>

Irrespective of robustness, a strong tendency for a positive sign of the coefficient of the number of family members in the wave before the move is observable.

4.3.6.1.3 Sensitivity for subgroups of dependent variables

4.3.6.1.3.1 Overview

<u>Robustness</u>

The fragility and non-significance of the number of family members in the wave before the move in explaining the probability of being risk-averse over all 4,536 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three finding are noteworthy: First, comparing results of subgroups referring to different risk-measures, I find that the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria in the subgroup referring to risk-measure semi-variances (LP2) are more than three times the percentages of robust relations in the subgroup referring to variance (Var). Second, comparing the results of subgroups referring to weighted versus unweighted samples, I find that the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria in the subgroup referring to unweighted samples (Unw) are more than three times the percentages of robust relations in the subgroup referring to weighted samples (Wei). Third, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (i.e, 2Kat3, 2Kat1, 4Kat2), (i) the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria decrease, and (ii) the strong tendency for a positive sign is less pronounced.

Tendency for a sign

The finding of a strong tendency for a positive sign of the number of family members in the wave before the move for all 4,536 dependent variables statistically investigated is not sensitive to any variation investigated in my study. In detail, the number of family members in the wave before the move is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.6.1.3.2 Detailed statistical reasoning and interpretation

Since the effect of the number of family members in the wave before the move on the probability of being risk-averse is not sensitive to any variation of competing solutions analyzed in my study, I do not discuss the various subgroups here. For detailed numerical results please refer to Table 43, p. 242.

4.3.6.2 Number of family members in the wave after the move

4.3.6.2.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The influence of the number of family members after the move on the probability of being risk-averse can be statistically investigated for all 4,536 dependent variables since (quasi-) complete separation does not occur. Based on this sample, the number of family members after the move does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Concerning significance, no statement can be made since Leamer's criterion of robustness does not find a robust relation of the number of family members in the wave after the move to any of the dependent variables. Yet, I derive a strong tendency for a negative sign of the number of family members after the move, which means the probability of being risk-averse is lower for Heads that live in a greater family after the move compared of Heads that live in a smaller family after the move.

Comparison to the literature

My findings are not in line with the majority of the literature as follows:[297] First, most authors find a significant effect of family size on risk-attitude in the non-migration domain, while only Harrison, Lau, and Rutström (2007) and Säve-Söderbergh (2012) do not find a significant effect. Second, the strong tendency for a negative sign found in my study contradicts previous findings in developed countries where Heads of bigger families showed significantly higher risk-aversion in the non-migration domain compared to Heads of smaller families. Third, a negative effect of family size on risk-attitudes in the non-migration domain was only found for less developed countries.

The difference of my findings compared to those of the literature is not surprising to me for two reasons. First, risk-attitudes are domain-specific and the domain of migration has not been surveyed yet. Second, the number of family members in the wave after the move might not fully capture the situation at the time of the move since the number of family members in the wave after the move does not necessarily relate to the number of family members that actually move together with Head.

[297] For a detailed discussion of findings on the relation for family size and risk-attitudes in the previous empirical literature please refer to Part C, p. 170.

4.3.6.2.2 Detailed statistical reasoning

All statistical results on the influence of the number of family members in the wave after the move on the probability of being risk-averse are summarized in Table 44, p. 247.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,530	0%	0.1%	0%	(-)	14%	45%		50%		6%	19%	(-)	48%	
Ind	3,882	0%	0.2%	0%	(-)	13%	42%		49%		6%	18%	(-)	47%	
Fam	648	0%	0.0%	n.a.		15%	66%		56%		8%	24%	(-)	53%	
Ann	1,510	0%	0.1%	0%	(-)	14%	46%		50%		6%	20%	(-)	48%	
Wor	1,510	0%	0.1%	0%	(-)	14%	45%		50%		6%	19%	(-)	48%	
Lif	1,510	0%	0.1%	0%	(-)	14%	46%		50%		6%	19%	(-)	48%	
Var	2,268	0%	0.0%	n.a.		11%	97%	(+)	68%		2%	93%	(+)	64%	
LP2	2,262	0%	0.3%	0%	(-)	17%	12%	(-)	33%		10%	6%	(-)	32%	
Ed1	1,293	0%	0.2%	0%	(-)	14%	50%		49%		7%	18%	(-)	45%	
Ed2	645	0%	0.5%	0%	(-)	15%	45%		43%		6%	17%	(-)	41%	
Ed3	648	0%	0.0%	n.a.		13%	43%		58%		6%	23%	(-)	54%	
Ed4	648	0%	0.0%	n.a.		13%	41%		58%		6%	23%	(-)	54%	
Ed5	648	0%	0.0%	n.a.		13%	42%		46%		6%	17%	(-)	46%	
Ed6	648	0%	0.0%	n.a.		14%	47%		48%		6%	17%	(-)	49%	
Wei	2,262	0%	0.3%	0%	(-)	11%	28%		59%		6%	30%		58%	
Unw	2,268	0%	0.0%	n.a.		16%	57%		42%		7%	10%	(-)	38%	
Sep	2,268	0%	0.0%	n.a.		18%	53%		51%		6%	13%	(-)	50%	
Poo	2,262	0%	0.3%	0%	(-)	10%	33%		49%		6%	25%	(-)	46%	
One	1,512	0%	0.0%	n.a.		4%	84%	(+)	58%		3%	88%	(+)	58%	
Ad1	1,512	0%	0.4%	0%	(-)	19%	40%		47%		6%	0%	(-)	43%	
Ad2	1,506	0%	0.0%	n.a.		18%	43%		46%		9%	11%	(-)	42%	
L1	1,512	0%	0.0%	n.a.		1%	33%		50%		1%	67%		52%	
L2	1,512	0%	0.0%	n.a.		10%	72%		47%		4%	0%	(-)	38%	
Ma	1,506	0%	0.4%	0%	(-)	30%	38%		54%		13%	23%	(-)	53%	
2Kat3	1,512	0%	0.0%	n.a.		15%	54%		58%		5%	59%		56%	
2Kat1	1,506	0%	0.4%	0%	(-)	14%	46%		53%		6%	6%	(-)	48%	
4Kat2	1,512	0%	0.0%	n.a.		13%	35%		40%		8%	4%	(-)	40%	

Table 44: Results of the extreme bounds analysis for the independent variable "number of family members in the wave after the move" aggregated over all 4,536 dependent variables and subgroups of dependent variables.
Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

<u>Sample size and (quasi-) complete separation</u>

(Quasi-) Complete separation does not occur. Hence, the statistical reasoning on robustness, sign, and significance of the number of family members in the wave after the move is based on all 4,536 dependent variables (Column (2), Table 44, p. 247).

<u>Robustness and significance based on Leamer's criterion</u>

Applying Leamer's criterion of robustness, the number of family members in the wave after the move is clearly classified as fragile, i.e., the number of family members in the wave after the move does not show a robust influence on any of the 4,536 dependent variables (Column (4), Table 44, p. 247). Since no robust relation is found, a statement on significance cannot be derived (Columns (5) and (6), Table 44, p. 247).

<u>Robustness based on Sala-I-Martin's criterion</u>

Applying Sala-I-Martin's criterion of robustness, the number of family members in the wave after the move is also clearly classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 14% (6%) of the 4,536 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 44, p. 247).

<u>Tendency for a sign based on Leamer's and Sala-I-Martin's criteria</u>

Irrespective of robustness, a strong tendency for a negative sign of the coefficient of the number of family members in the wave after the move is observable. Note that the percentage of positive relations found by Leamer's criterion cannot be considered because Leamer's criterion does not find any robust relations.

4.3.6.2.3 Sensitivity for subgroups of dependent variables

4.3.6.2.3.1 Overview

<u>Robustness</u>

The fragility and non-significance of the number of family members in the wave after the move in explaining the probability of being risk-averse over all 4,536 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing results of subgroups referring to different measurements of predictive errors, I find that in the subgroup referring to minimizing the maximum predictive error (Ma) the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion is about 30 times (13 times) the percentage in the subgroup referring to minimizing the sum of predictive errors (L1), and three times (three times) the percentage of robust relations in the subgroup referring to minimizing the sum of squared predictive errors (L2).

Tendency for a sign

The finding of a strong tendency for a negative sign of the number of family members in the wave after the move for all 4,536 dependent variables statistically investigated is not sensitive to (i) variations of the planning period (denoted by Ann, Wor, and Lif), (ii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), and (iii) not sensitive to variations in the measurement of predictive errors (denoted by L1, L2, and Ma).

However, the finding of a strong tendency for a negative sign is sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) different risk-measures (denoted by Var and LP2), (iii) variations of the education definition (denoted by Ed1 to Ed6), (iv) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (v) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), and (vi) sensitive to different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.6.2.3.2 Detailed statistical reasoning and interpretation

Two facts are identical for all subgroups and will, therefore, not be mentioned again for every subgroup. They are, first, (quasi-)complete separation does not occur for the total of all 4,536 dependent variables and hence is of no concern for any subgroups of dependent variables. Second, the relation between the number of family members in the wave after the move and the probability of being risk-averse is not robust for any subgroup by Leamer's criterion of robustness. Consequently, only robustness by Sala-I-Martin's criterion and the tendencies for a sign are discussed in this section.

Detailed comparison of subgroups of dependent variables relating to different decision problems

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is sensitive to whether the migration decision is estimated based on **Head's individual income or family income** (denoted by Ind and Fam). First, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in both subgroups.

Second, I observe a strong tendency for a negative sign in the subgroup referring to Head's individual income (Ind), and no tendency for a sign in the subgroup referring to family income (Fam).

I interpret these findings as follows: The probability of being risk-averse towards variations of Head's individual income (Ind) seems to be the smaller, the greater the number of family members after the move. At the same time Heads with a greater number of family members after the move do not seem to systematically differ in their risk-attitude towards variation of the family income (Fam). A first explanation for this finding might be the considerably different sample sizes in the two subgroups. Since migration decisions based on family income (Fam) are only estimated applying education definition one, the number of dependent variables in the subgroup referring to family income (Fam) amounts to one sixth of the dependent variables in the subgroup referring to Head's individual income (Ind). A second explanation might be that the probability of Head being the only one who contributes to family income decreases when family size increases. Hence, family members of larger families are less dependent on Head's individual income and therefore less probable of being risk-averse towards variations of Head's individual income (Ind). Once variations of family income (Fam) is taken into account (rather than only Head's income), family size does not matter anymore regarding the probability of being risk-averse.

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is not sensitive to **different planning periods** investigated in my study, i.e., one year, time until reaching full retirement age, and time until reaching life expectancy (denoted by Ann, Wor, and Lif). For all subgroups I find that, first, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion, and second, a strong tendency for a negative sign can be observed.

  I interpret these findings as being due to the way income parameters for longer planning periods have been estimated (for a detailed discussion of this argument see the corresponding interpretation for Male, p. 204).

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is sensitive to **different risk-measures** variance and semi-variance (denoted by Var and LP2). First, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in both subgroups. Second, I observe a strong tendency for a positive sign in the subgroup referring to variance (Var), and a strong tendency for a negative sign in the subgroup referring to semi-variance (LP2).

I interpreted the findings as follows: Since the risk-measure variance fully includes downside-risk, the combination of tendencies for the two risk-measures variance and semi-variance can only be true if Head's of larger families after the move consider deviations of income above the expected value as more problematic compared to Head's of smaller families in the wave after the move than deviations below the expected value. A possible explanation for this finding might be that larger families have a better downside protection compared to smaller families, e.g., larger families might have more people who contribute to family income.

<u>Detailed comparison of subgroups of dependent variables relating to different ways to gain data input</u>

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is sensitive to **different education definitions** (denoted by Ed1 to Ed6). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income only education definition one was applied. Therefore, the number of dependent variables to be analyzed amounts to 648 for education definitions two to six, while education definition one additionally includes all migration decisions based on family income.

  Concerning results, it is found: First, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in all subgroups. Second, I observe a strong tendency for a negative sign in the subgroup referring to education definitions one, two, four, and six (Ed1, Ed2, Ed5, Ed6), and no tendency for a sign in the subgroup referring to education definitions three and four (Ed3, Ed4).

  To be able to interpret the findings of pairwise similar subgroups, recall that the six education definitions were derived from two competing solutions to each of the two following questions: (1) The question on how to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File, and (2) whether education variables from years before or after the move should be considered first when education variables are not available in each year.

  Therefore, I interpreted my findings as follows: First, subgroups Edu3 and Edu4 that show different tendency for a sign than all others subgroups differ in their answer to Question (1). Edu3 and Edu4 always prefer education variables from the Single-Year Family Files, while the other subgroups Ed1, Ed2, Ed5, and Ed6 always prefer education variables from the Cross-Year Individual File. Therefore, I conclude that the answer to Question (1) is decisive for the sensitivity found. Second, Ed1, Ed3, and Ed5 answer Question (2) by searching for education information in the years centered around the moving date, while subgroups Ed2, Ed4, and Ed6 first search in the years before the move and then in years after the move. Obviously, the different answers to Question (2) do not alter the result of how the number of family members in the wave after the move is related to the probability of being risk-averse in the migration context.

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is not sensitive to whether **income parameters are estimated based on weighted or unweighted samples** (denoted by Wei and Unw). For both subgroups I find that, first, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion, and second, no tendency for a sign can be observed.

  I interpret these findings as follows: Obviously, the concern of biased results due to biased estimations of the income parameters for weighted samples cannot be confirmed for the relation of the number of family members after the move to the probability of being risk-averse in my study (for a detailed discussion of this argument see the corresponding interpretation for Male, p. 207).

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is sensitive to the **type of clustering** of people from which income parameters are estimated, i.e., separate clustering for each year versus pooled clustering with same clusters in all years (denoted by Sep and Poo). First, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in both subgroups. Second, I observe no tendency for a sign in the subgroup referring to separate pooling (Sep), and a vague tendency for a negative sign in the subgroup referring to pooled clustering (Poo).

  I interpret these findings as follows: The considerably different age-clusters resulting from the two types of clustering seem to be decisive for the effect of the number of family members after the move on the probability of being risk-averse. Still, this difference should not be overrated since the subgroup referring to separate clustering (Sep) only fails to meet the criteria for vague tendency for a negative sign by 3% points.

- The relation of the number of family members in the wave after the move to the probability of being risk-averse is sensitive to different **time periods from which income parameters are**

estimated, namely annual income data of one single year, annual income date from three years around the moving date adjusted for inflation, and annual income data from three years around the moving date not adjusted for inflation (denoted by One, Ad1, and Ad2). First, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in all subgroups. Second, I find a strong tendency for a positive sign in the subgroup referring to one year's data (One), while a strong tendency for a negative sign can be observed for the two subgroups referring to three year's data (Ad1 and Ad2).

I interpret these findings as follows: It seems as if the results of the number of family members after the move and its relation to the probability of being risk-averse are sensitive to whether estimations are based on one year's data compared to three year's of data, but not sensitive to whether the three year's data are adjusted for inflation or not. A possible explanation might be that income parameters systematically differs for different years which results in systematically different income parameters when they are estimated from one year's data (One) compared to three year's data (Ad1 and Ad2).

Detailed comparison of subgroups of dependent variables relating to different estimation procedures of risk-attitudes

The relation of the number of family members in the wave after the move to the probability of being risk-averse is not sensitive to **different measurements of predictive errors**, namely L_p-norms one-, two- and infinity-norms where $p = \{1,2,\infty\}$ (denoted by L1, L2, and Ma). First, although the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in all subgroups, it is noteworthy that in the subgroup referring to minimizing the maximum predictive errors (Ma) the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion is about 30 times (13 times) the percentage in the subgroup referring to minimizing the sum of predictive errors (L1), and three times (three times) the percentage of robust relations in the subgroup referring to minimizing the sum of squared predictive errors (L2). Second, no tendency for a sign can be observed.

I interpret these findings as follows: It seems as if the maximum predictive error is decisive since the subgroups that additionally account for smaller predictive errors (L1 and L2) exhibit qualitative the same results. This argument might also explain why the subgroup referring to minimizing the maximum error (Ma) is also the subgroup with the highest percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria.

Detailed comparison of subgroups of dependent variables relating to different transformation rules to code the binary dependent variables

The relation of the number of family members in the wave after the move to the probability of being risk-averse is sensitive to **different transformation rules** (denoted by 2Kat1, 2Kat3, and 4Kat2). First, the number of family members after the move is classified as fragile by Sala-I-Martin's criterion in all subgroups. Second, I find a vague tendency for a positive sign for the subgroup referring to all migrants (2Kat3), no tendency for a sign for the subgroup that includes only the 95% migrants exhibiting the strongest degree of risk-attitude (2Kat1), and a strong tendency for a negative sign for the subgroup including only the 50% migrants exhibiting the strongest degree of risk-attitude (4Kat2).

I interpret these findings as follows: When migrants with more pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the tendency for a sign turns more negative. This could be interpreted as confirming the initial idea of the different transformation rules. The idea was that biased results due to false classification of migrants into risk-averse and risk-seeking by applying a threshold value of zero could be detected when results of the different transformation rules were compared. This could actually be true since results for the most extreme transformation rule (4Kat2) are in line with result from the total of all 4,536 dependent variables discussed in Part C, Section 4.3.6.2.1.

4.3.6.3 Conclusion on the specification of family size

The strong tendency for a positive sign for the variable of the number of family members in the wave before the move (Part C, Section 4.3.6.1) contradicts the finding of strong tendency for a negative sign for the variable on the number of family members in the wave after the move (Part C, Section 4.3.6.2). Obviously, both variables do not capture the same effect as was suggested in Part C, Section 4.1.1.

4.3.7 Single-, pair-, and family-moves and their relation to risk-attitudes

The previous section has shown that the number of family members before/after the move might not be appropriate in the context of migration since it does not relate to the number of people actually moving together. Alternative indicators might be (i) the number of people actually moving with Head and (ii) their personal relation to Head. Both aspects are captured by dummy variables Single, Pair, and Family which are discussed in detail in this section.

4.3.7.1 Single-moves

4.3.7.1.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

In my study, the dummy variable Single indicates whether Head is moving on his own or not (see Part B, Section 2.4.4 for a detailed definition of this variable). The influence of Single on the probability of being risk-averse can be statistically investigated for 4,276 dependent variables since quasi-complete separation only occurs for about 6% of the 4,536 dependent variables. Based on this sample, Single does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Concerning significance of Single, no statement can be made since Leamer's criterion of robustness does not find a robust relation of Single to any of the dependent variables. Yet, I derive a strong tendency for a positive sign of Single.

Comparison to the literature and interpretation

Since the effect of moving alone as captured by the variable Single has not been surveyed yet, my findings are new to the literature.

The strong tendency for a positive sign of Single indicates that Heads moving on their own have a higher probability of being risk-averse compared to Heads that move together with other people. A possible explanation might be that Heads moving on their own rely on themselves and have nobody to support them financially. In contrast, people moving together as a family might have more people that contribute to family income and support each other. Put differently, they have the possibility to diversify their migration-risk.

4.3.7.1.2 Detailed statistical reasoning

All statistical results on the influence of moving alone (as indicated by Single) on the probability of being risk-averse are summarized in Table 45, p. 256.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,276	6%	0%	n.a.		41%	93%	(+)	88%	(+)	20%	94%	(+)	88%	(+)
Ind	3,702	5%	0%	n.a.		40%	93%	(+)	88%	(+)	19%	93%	(+)	89%	(+)
Fam	574	11%	0%	n.a.		48%	94%	(+)	89%	(+)	27%	97%	(+)	87%	(+)
Ann	1,425	6%	0%	n.a.		41%	93%	(+)	89%	(+)	20%	95%	(+)	88%	(+)
Wor	1,426	6%	0%	n.a.		41%	92%	(+)	89%	(+)	20%	93%	(+)	88%	(+)
Lif	1,425	6%	0%	n.a.		41%	93%	(+)	88%	(+)	20%	93%	(+)	88%	(+)
Var	2,268	0%	0%	n.a.		33%	89%	(+)	89%	(+)	6%	89%	(+)	87%	(+)
LP2	2,008	11%	0%	n.a.		50%	96%	(+)	88%	(+)	36%	95%	(+)	90%	(+)
Ed1	1,183	9%	0%	n.a.		45%	92%	(+)	89%	(+)	24%	94%	(+)	88%	(+)
Ed2	609	6%	0%	n.a.		41%	94%	(+)	89%	(+)	22%	94%	(+)	90%	(+)
Ed3	612	6%	0%	n.a.		43%	97%	(+)	90%	(+)	20%	95%	(+)	89%	(+)
Ed4	612	6%	0%	n.a.		35%	94%	(+)	88%	(+)	17%	94%	(+)	88%	(+)
Ed5	630	3%	0%	n.a.		37%	88%	(+)	88%	(+)	15%	91%	(+)	88%	(+)
Ed6	630	3%	0%	n.a.		40%	91%	(+)	87%	(+)	17%	91%	(+)	89%	(+)
Wei	2,201	3%	0%	n.a.		30%	87%	(+)	85%	(+)	15%	88%	(+)	85%	(+)
Unw	2,075	9%	0%	n.a.		52%	96%	(+)	92%	(+)	24%	97%	(+)	92%	(+)
Sep	2,075	9%	0%	n.a.		34%	92%	(+)	88%	(+)	15%	99%	(+)	88%	(+)
Poo	2,201	3%	0%	n.a.		47%	93%	(+)	89%	(+)	24%	91%	(+)	89%	(+)
One	1,505	0%	0%	n.a.		36%	86%	(+)	80%	(+)	8%	97%	(+)	79%	(+)
Ad1	1,389	8%	0%	n.a.		44%	94%	(+)	93%	(+)	26%	92%	(+)	93%	(+)
Ad2	1,382	9%	0%	n.a.		42%	97%	(+)	94%	(+)	26%	95%	(+)	94%	(+)
L1	1,512	0%	0%	n.a.		34%	93%	(+)	92%	(+)	14%	100%	(+)	90%	(+)
L2	1,394	8%	0%	n.a.		59%	99%	(+)	98%	(+)	31%	100%	(+)	99%	(+)
Ma	1,370	9%	0%	n.a.		30%	80%	(+)	75%	(+)	14%	71%		76%	(+)
2Kat3	1,488	2%	0%	n.a.		41%	97%	(+)	91%	(+)	14%	100%	(+)	90%	(+)
2Kat1	1,422	6%	0%	n.a.		39%	95%	(+)	88%	(+)	18%	93%	(+)	88%	(+)
4Kat2	1,366	10%	0%	n.a.		43%	87%	(+)	87%	(+)	27%	91%	(+)	87%	(+)

Table 45: Results of the extreme bounds analysis for the independent variable Single aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Sample size and (quasi-) complete separation

Since quasi-complete separation only occurs for about 6% of the 4,536 dependent variables (Column (3), Table 45, p. 256), the statistical reasoning on robustness, sign, and significance of Single is based on a sufficiently high sample of 4,276 dependent variables (Column (2), Table 45, p. 256).

Concerning the 260 dependent variables (about 6% of 4,536, Column (3), Table 45, p. 256) where quasi-complete separation occurs, it is noteworthy that all single-movers in the sample are risk-averse.

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, Single is clearly classified as fragile, i.e., Single does not show a robust influence on any of the 4,276 dependent variables (Column (4), Table 45, p. 256). Since no robust relation is found, a statement on significance cannot be made (Columns (5) and (6), Table 45, p. 256).

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, Single is also clearly classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 41% (20%) of the 4,276 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 45, p. 256).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, a strong tendency for a positive sign of the coefficient of Single is observed.

4.3.7.1.3 Sensitivity for subgroups of dependent variables

4.3.7.1.3.1 Overview

Robustness

The fragility and non-significance of Single in explaining the probability of being risk-averse over the 4,276 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three findings are noteworthy: First, comparing results of subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variances (LP2) is about 1.5 times (6 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria in the subgroup referring to unweighted samples (Unw) are about 1.6 times the percentages of robust relations in the subgroup referring to weighted samples (Wei). Third, comparing results of subgroups referring to different transformation rules, I find that when only migrants with the most pronounced degrees of risk-attitude are included in the sample (4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria are highest compared to subgroups that include also migrants with less pronounced degrees of risk-attitude (2Kat3 and 2Kat1).

Tendency for a sign

The finding of a strong tendency for a positive sign of Single for the 4,276 dependent variables statistically investigated is not sensitive to any variation investigated in my study. In detail, the tendency for a sign of Single is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.7.1.3.2 Detailed statistical reasoning and interpretation

Since the effect of Single on the probability of being risk-averse is not sensitive to any variation of competing solutions analyzed in my study, I do not discuss the various subgroups here. For detailed numerical results please refer to Table 45, p. 256.

4.3.7.2 Pair-moves

4.3.7.2.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

In my study, the dummy variable Pair indicates whether Head is moving together with his partner and possibly newborns (see Part B, Section 2.4.4 for a detailed definition of this variable). The influence of Pair on the probability of being risk-averse can be statistically investigated for 4,387 dependent variables since quasi-complete separation only occurs for about 3% of the 4,536 dependent variables. Based on this sample, Pair does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Consequently, Pair is also not significantly related to the probability of being risk-averse in the migration context. Moreover, I observe no tendency for sign of Pair.

Comparison to the literature and interpretation

Since the effect of being a pair and moving together (possibly with newborns) as captured by the variable Pair has not been surveyed yet, my findings are new to the literature.

The non-robustness and non-existence of a tendency for a sign on Pair indicates that Heads moving with their partner (and possibly newborns) do not have a higher or lower probability of being risk-averse in the migration context than all other migrants. A possible explanation for my result might be found in the composition of the reference category of Pair. The reference category comprises two, potentially heterogeneous, groups (single- and family-movers). Assume, for example, that single- and family-movers show different tendencies for a sign and pair-movers are different to only single-movers, but not family-movers. Then Pair might not exhibit a statistically observable tendency for a sign although pair-movers might be different in their probability of being risk-averse compared to single-movers.

4.3.7.2.2 Detailed statistical reasoning

All statistical results on the influence of Heads moving with their partner (and possibly newborns) (as indicated by Pair) on the probability of being risk-averse are summarized in Table 46, p. 260.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,387	3%	0.5%	0%	(-)	27%	71%		47%		15%	58%		49%	
Ind	3,742	4%	0.6%	0%	(-)	28%	74%		49%		15%	63%		51%	
Fam	645	0%	0.2%	0%	(-)	25%	51%		36%		16%	27%		38%	
Ann	1,465	3%	0.5%	0%	(-)	28%	72%		47%		16%	59%		49%	
Wor	1,461	3%	0.5%	0%	(-)	27%	69%		48%		15%	56%		50%	
Lif	1,461	3%	0.6%	0%	(-)	28%	71%		48%		16%	58%		49%	
Var	2,233	2%	0.0%	n.a.		25%	82%	(+)	48%		15%	97%	(+)	51%	
LP2	2,154	5%	1.1%	0%	(-)	30%	60%		46%		16%	20%	(-)	47%	
Ed1	1,270	2%	0.8%	0%	(-)	26%	61%		42%		16%	38%		44%	
Ed2	625	4%	1.9%	0%	(-)	29%	69%		47%		17%	51%		50%	
Ed3	621	4%	0.0%	n.a.		25%	78%	(+)	53%		14%	79%	(+)	53%	
Ed4	621	4%	0.0%	n.a.		29%	72%		52%		15%	78%	(+)	52%	
Ed5	625	4%	0.0%	n.a.		28%	76%	(+)	48%		15%	61%		50%	
Ed6	625	4%	0.3%	0%	(-)	29%	77%	(+)	48%		16%	63%		50%	
Wei	2,154	5%	0.5%	0%	(-)	34%	80%	(+)	57%		18%	79%	(+)	59%	
Unw	2,233	2%	0.5%	0%	(-)	21%	55%		38%		13%	28%		40%	
Sep	2,220	2%	0.5%	0%	(-)	27%	65%		43%		17%	44%		42%	
Poo	2,167	4%	0.5%	0%	(-)	28%	77%	(+)	52%		14%	75%	(+)	56%	
One	1,512	0%	0.1%	0%	(-)	34%	74%		48%		6%	78%	(+)	52%	
Ad1	1,441	5%	0.6%	0%	(-)	25%	72%		48%		21%	58%		48%	
Ad2	1,434	5%	0.9%	0%	(-)	23%	64%		46%		20%	51%		47%	
L1	1,512	0%	0.0%	n.a.		25%	94%	(+)	55%		16%	85%	(+)	60%	
L2	1,512	0%	0.0%	n.a.		32%	58%		40%		11%	34%		37%	
Ma	1,363	10%	1.6%	0%	(-)	25%	63%		47%		19%	47%		50%	
2Kat3	1,512	0%	0.8%	0%	(-)	17%	67%		34%		12%	52%		37%	
2Kat1	1,470	3%	0.0%	n.a.		39%	81%	(+)	55%		19%	76%	(+)	55%	
4Kat2	1,405	7%	0.8%	0%	(-)	27%	57%		55%		15%	39%		55%	

Table 46: Results of the extreme bounds analysis for the independent variable Pair aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

<u>Sample size and (quasi-) complete separation</u>

Since quasi-complete separation only occurs for about 3% of the 4,536 dependent variables (Column (3), Table 46, p. 260), the statistical reasoning on robustness, sign, and significance of Pair is based on a sufficiently high sample of 4,387 dependent variables (Column (2), Table 46, p. 260).

Concerning the 149 dependent variables (about 3% of 4,536, Column (3), Table 46, p. 260) where quasi-complete separation occurs, it is noteworthy that all pair-movers in the sample are risk-averse.

<u>Robustness and significance based on Leamer's criterion</u>

Applying Leamer's criterion of robustness, Pair is clearly classified as fragile, i.e., Pair does only show a robust significant influence on 0.5% of the 4,387 dependent variables analyzed (Column (4), Table 46, p. 260). Consequently, Pair is classified as not significantly related to probability of being risk-averse in the migration context.

<u>Robustness based on Sala-I-Martin's criterion</u>

Applying Sala-I-Martin's criterion of robustness, Pair is also clearly classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 27% (15%) of the 4,387 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 46, p. 260).

<u>Tendency for a sign based on Leamer's and Sala-I-Martin's criteria</u>

Irrespective of robustness, no tendency for a sign of the coefficient of Pair is observable.

4.3.7.2.3 Sensitivity for subgroups of dependent variables

4.3.7.2.3.1 Overview

<u>Robustness</u>

The fragility and non-significance of Pair in explaining the probability of being risk-averse over the 4,387 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, one finding is noteworthy: Comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to weighted samples (Wei) is about 1.6

times (1.4 times) the percentage of robust relations in the subgroup referring to unweighted samples (Unw).

Tendency for a sign

The finding of no tendency for a sign of Pair for the 4,387 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the risk-measure (denoted by Var and LP2), and not sensitive to (iv) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2).

However, the finding of no tendency for a sign of Pair is sensitive to (i) variations of the education definition (denoted by Ed1 to Ed6), (ii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iii) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (iv) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and sensitive to (v) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.7.2.3.2 Detailed statistical reasoning and interpretation

Two facts are identical for all subgroups and will therefore not be mentioned again for every subgroup. They are, first, quasi-complete separation is of no concern for any subgroup since it does not exceed 10% of the dependent variables in any subgroup. Second, the relation between Pair and the probability of being risk-averse is not robust by any criterion in any subgroup where percentages of robust relations are equally low in all subgroups. Hence, third, Pair is also not significantly related to the probability of being risk-averse in any subgroup. Consequently, only the sensitivity of the tendency for a sign on Pair is discussed in this section.

Detailed comparison of subgroups of dependent variables relating to different decision problems

- The relation of Pair to the probability of being risk-averse is not sensitive to whether the migration decision is estimated based on **Head's individual income or family income** (denoted by Ind and Fam), i.e., both subgroups exhibit no tendency for a sign. This is especially noteworthy since the number of dependent variables analyzed in both subgroups differs considerably. For migration decisions based on family income only education definition one was applied.

Therefore, the number of dependent variables in the subgroup referring to family income (Fam) amounts to one sixth of the dependent variables based in Head's individual income (Ind).

I interpret these findings as follows: Irrespective of whether Head's individual income or family income is considered, Heads moving with their partner (or possibly newborns) do not differ in their probability of being risk-averse in the migration context compared to the total of single-movers and family-movers.

- The relation of Pair to the probability of being risk-averse is not sensitive to **different planning periods** investigated in my study, i.e., one year, time until reaching full retirement age, and time until reaching life expectancy (denoted by Ann, Wor, and Lif), i.e., both subgroups exhibit no tendency for a sign.

  I interpret these findings as being due to the way income parameters for longer planning periods have been estimated (for a detailed discussion of this argument see the corresponding interpretation for Male, p. 204).

- The relation of Pair to the probability of being risk-averse is not sensitive to **different risk-measures** variance and semi-variance (denoted by Var and LP2), i.e., both subgroups exhibit no tendency for a sign.

  I interpreted the findings as follows: Heads moving with their partner (and possibly newborns) do not differ in their risk-attitude towards either variance (Var) or semi-variance (LP2) compared to the total of single-movers and family-movers.

Detailed comparison of subgroups of dependent variables relating to different ways to gain data input

- The relation of Pair to the probability of being risk-averse is sensitive to **different education definitions** (denoted by Ed1 to Ed6). Concerning the number of dependent variables analyzed, it is noteworthy that for migration decisions based on family income, only education definition one was applied. Therefore, the number of dependent variables to be analyzed amounts to 648 for education definitions two to six, while education definition one additionally includes all migration decisions based on family income. Concerning the tendency for a sign, I observe no tendency for a sign in the subgroups referring to education definitions one, two, four, and six (Ed1, Ed2, Ed5, Ed6), but a strong tendency for a positive sign in the subgroup referring to education definitions three and four (Ed3, Ed4).

To be able to interpret the finding of pair wise similar subgroups, recall that the six education definitions were derived from two competing solutions to each of the two following questions: (1) The question on how to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File, and (2) whether education variables from years before or after the move should be considered first when education variables are not available in each year.

Therefore, I interpreted my findings as follows: First, subgroups Edu3 and Edu4 that show different tendencies for a sign than all others subgroups differ in their answer to Question (1). Edu3 and Edu4 always prefer education variables from the Single-Year Family Files, while the other subgroups Ed1, Ed2, Ed5, and Ed6 always prefer education variables from the Cross-Year Individual File. Therefore, I conclude that the answer to Question (1) is decisive for the sensitivity found. Second, Ed1, Ed3, and Ed5 answer Question (2) by searching for education information in the years centered around the moving date, while subgroups Ed2, Ed4, and Ed6 first search in the years before the move and then in years after the move. Obviously, the different answers to Question (2) do not alter the result of how Pair is related to the probability of being risk-averse in the migration context.

- The relation of Pair to the probability of being risk-averse is sensitive to whether the **income parameters are estimated based on weighted or unweighted samples** (denoted by Wei and Unw). For the subgroup referring to weighted samples (Wei) I observe a strong tendency for a positive sign, but I observe no tendency for a sign in the subgroup referring to unweighted samples (Unw).

  I interpret these findings as follows: Obviously, it makes a difference for the influence of Pair on the probability of being risk-averse whether samples are weighted or not. Since both approaches are afflicted with problems, it is not clear whether it is (i) the unrepresentative sample of the subgroup referring to unweighted samples (Wei) or (ii) the weighting of the sample (Wei) that causes biased estimates of income parameters.

- The relation of Pair to the probability of being risk-averse is sensitive to the **type of clustering** of people from which income parameters are estimated, i.e., separate clustering for each year versus pooled clustering with same clusters in all years (denoted by Sep and Poo). For the subgroup referring to separate clustering in each year (Sep), I observe no tendency for a sign, but in the subgroup referring to pooled clustering (Poo) I observe a strong tendency for a positive sign.

I interpret these findings as follows: The considerably different age-clusters resulting from the two types of clustering seem to be decisive for the effect of Pair on the probability of being risk-averse. Put differently, when migrants are defined into different age-clusters from which income parameters are estimated, the effect of Pair on the probability of being risk-averse seems to differ.

- The relation of Pair to the probability of being risk-averse is not sensitive to different **time periods from which income parameters are estimated**, namely annual income data of one single year, annual income date from three years around the moving date adjusted for inflation, and annual income data from three years around the moving date not adjusted for inflation (denoted by One, Ad1, and Ad2). That is, all subgroups exhibit no tendency for a sign.

  I interpret these findings as follows: It seems as if the results of Pair and its relation to the probability of being risk-averse are not sensitive to whether estimations are based on one year's data compared to three year's of data, and also not sensitive to whether the three year's data are adjusted for inflation or not.

<u>Detailed comparison of subgroups of dependent variables relating to different estimation procedures of risk-attitudes</u>

The relation of Pair to the probability of being risk-averse is sensitive to **different measurements of predictive errors**, namely L_p-norms one-, two- and infinity-norms where $p = \{1,2,\infty\}$ (denoted by L1, L2, and Ma). For the subgroup referring to minimizing the sum of predictive errors (L1) I observe a strong tendency for a positive sign. In contrast, in the subgroups referring to minimizing the sum of squared predictive errors (L2) and in the subgroups referring to minimizing the maximum predictive error (Ma) I observe no tendency for a sign.

I interpret these findings as follows: It seems as if a great number of smaller predictive errors exists that is responsible for the difference in the tendencies for a sign observable in the subgroup referring to minimizing the sum of predictive errors (L1) compared to subgroups that put greater weight to greater predictive errors (L2 and Ma).

Detailed comparison of subgroups of dependent variables relating to different transformation rules to code the binary dependent variables

The relation of Pair to the probability of being risk-averse is sensitive to **different transformation rules** (denoted by 2Kat1, 2Kat3, and 4Kat2). For the subgroups referring to all migrants (2Kat3) and the subgroup referring to only those migrants with the most pronounced degree of risk-attitude (4Kat2) I observe no tendency for sign. In contrast, in the subgroup referring to the transformation rule that deletes only those 5% of migrants that exhibit the least pronounced degree of risk-attitude I observe a strong tendency for a positive sign.

I interpret these findings as follows: Concerning the effect of Pair on the probability of being risk-averse, it seems as if migrants with most pronounced degrees of risk-attitude (4Kat2) are similar to the total of all migrants (2Kat3). Further interpretations are impossible because the different transformations rules mix relative and absolute threshold values to delete migrants from the sample. This means, those migrants deleted from the sample under transformation rule 2Kat1 because they belong to those 5% with the least pronounced degree of risk-attitude are not necessarily the same migrants that are deleted from the sample when the most extreme transformation rule (4Kat2) is applied.

4.3.7.3 Family-moves

4.3.7.3.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

In my study, the dummy variable Family indicates whether Head is moving together with his family, i.e., either (i) together with his partner and further family members that are no newborns or (ii) without a partner but at least one other family member (see Part B, Section 2.4.4 for a detailed definition of this variable). The influence of Family on the probability of being risk-averse can be statistically investigated for 4,530 dependent variables since quasi-complete separation only occurs for about 0.1% of the 4,536 dependent variables. Based on this sample, Family does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Concerning significance of Family, no statement can be made since Leamer's criterion of robustness does not find a robust relation of Family to any of the dependent variables. Yet, I derive a strong tendency for a negative sign of Family.

Comparison to the literature and interpretation

Since the effect of moving as a family as captured by the variable Family has not been surveyed yet, my findings are new to the literature.

The strong tendency for a negative sign of Family indicates that Heads moving together with their family have a lower probability of being risk-averse than all other migrants. I can think of two explanations for this finding: First, when more than two people move together as a family, it might be that more than two people contribute to family income. This means, family members are able to diversify income risks. Since family income in my study is restricted to income of the two main earners, it might be that when more people contribute to family income, Heads are more willing to take risk concerning income of the two main earners. Second, irrespective of financial aspects, it might be that family members support each other psychologically which makes Heads more probable to take migration-risk compared to the Heads that move on their own (single-movers) or with their partner and possibly newborns (pair-move). Both explanations hint in the same direction and might reinforce each other.

4.3.7.3.2 Detailed statistical reasoning

All statistical results on the influence of moving together with a family (as indicated by Family) on the probability of being risk-averse are summarized in Table 47, p. 268.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,530	0.1%	0%	n.a.		25%	12%	(-)	23%	(-)	12%	7%	(-)	22%	(-)
Ind	3,882	0.2%	0%	n.a.		24%	12%	(-)	24%	(-)	11%	7%	(-)	22%	(-)
Fam	648	0.0%	0%	n.a.		28%	12%	(-)	20%	(-)	16%	8%	(-)	22%	(-)
Ann	1,510	0.1%	0%	n.a.		24%	13%	(-)	23%	(-)	12%	7%	(-)	22%	(-)
Wor	1,510	0.1%	0%	n.a.		25%	11%	(-)	23%	(-)	12%	7%	(-)	21%	(-)
Lif	1,510	0.1%	0%	n.a.		25%	12%	(-)	23%	(-)	12%	8%	(-)	22%	(-)
Var	2,268	0.0%	0%	n.a.		22%	19%	(-)	18%	(-)	4%	30%		19%	(-)
LP2	2,262	0.3%	0%	n.a.		27%	7%	(-)	28%		20%	3%	(-)	24%	(-)
Ed1	1,293	0.2%	0%	n.a.		27%	12%	(-)	23%	(-)	15%	8%	(-)	22%	(-)
Ed2	645	0.5%	0%	n.a.		27%	12%	(-)	24%	(-)	14%	9%	(-)	20%	(-)
Ed3	648	0.0%	0%	n.a.		27%	1%	(-)	20%	(-)	17%	0%	(-)	18%	(-)
Ed4	648	0.0%	0%	n.a.		23%	9%	(-)	24%	(-)	13%	7%	(-)	21%	(-)
Ed5	648	0.0%	0%	n.a.		18%	25%	(-)	25%	(-)	6%	16%	(-)	26%	
Ed6	648	0.0%	0%	n.a.		24%	18%	(-)	24%	(-)	5%	15%	(-)	25%	(-)
Wei	2,262	0.3%	0%	n.a.		22%	15%	(-)	32%		8%	11%	(-)	29%	
Unw	2,268	0.0%	0%	n.a.		27%	10%	(-)	15%	(-)	16%	5%	(-)	15%	(-)
Sep	2,268	0.0%	0%	n.a.		22%	10%	(-)	28%		11%	3%	(-)	24%	(-)
Poo	2,262	0.3%	0%	n.a.		27%	13%	(-)	18%	(-)	13%	11%	(-)	19%	(-)
One	1,512	0.0%	0%	n.a.		13%	52%		44%		6%	39%		44%	
Ad1	1,512	0.0%	0%	n.a.		31%	4%	(-)	13%	(-)	15%	3%	(-)	11%	(-)
Ad2	1,506	0.4%	0%	n.a.		29%	2%	(-)	13%	(-)	16%	0%	(-)	10%	(-)
L1	1,512	0.0%	0%	n.a.		20%	11%	(-)	19%	(-)	7%	10%	(-)	22%	(-)
L2	1,512	0.0%	0%	n.a.		32%	4%	(-)	10%	(-)	18%	1%	(-)	10%	(-)
Ma	1,506	0.4%	0%	n.a.		21%	25%	(-)	40%		11%	16%	(-)	34%	
2Kat3	1,512	0.0%	0%	n.a.		17%	5%	(-)	24%	(-)	3%	0%	(-)	23%	(-)
2Kat1	1,506	0.4%	0%	n.a.		21%	11%	(-)	25%	(-)	12%	11%	(-)	25%	(-)
4Kat2	1,512	0.0%	0%	n.a.		35%	16%	(-)	20%	(-)	21%	7%	(-)	18%	(-)

Table 47: Results of the extreme bounds analysis for the independent variable Family aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Sample size and (quasi-) complete separation

Since quasi-complete separation only occurs for about 0.1% of the 4,536 dependent variables (Column (3), Table 47, p. 268), the statistical reasoning on robustness, sign, and significance of Family is based on a sufficiently high sample of 4,530 dependent variables (Column (2), Table 47, p. 268).

Concerning the 6 dependent variables (about 0.1% of 4,536, Column (3), Table 47, p. 268) where quasi-complete separation occurs, it is noteworthy that all family-movers in the sample are risk-averse.

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, Family is clearly classified as fragile, i.e., Family does not show a robust influence on any of the 4,530 dependent variables (Column (4), Table 47, p. 268). Since no robust relation is found, a statement on significance cannot be made (Columns (5) and (6), Table 47, p. 268).

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, Family is also clearly classified as fragile. Sala-I-Martin's weighted (unweighted) criterion finds only 25% (12%) of the 4,530 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 47, p. 268).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, a strong tendency for a negative sign of the coefficient of Family is observable.

4.3.7.3.3 Sensitivity for subgroups of dependent variables

4.3.7.3.3.1 Overview

Robustness

The fragility and non-significance of Family in explaining the probability of being risk-averse over the 4,530 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, one finding is noteworthy: Comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are

deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a strong tendency for a negative sign of Family for the 4,530 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the risk-measure (denoted by Var and LP2), (iv) variations of the education definition (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (vii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and (viii) not sensitive to different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

However, the finding of a strong tendency for a negative sign of Family on the probability of being risk-averse for the total of all dependent variables is sensitive to whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2).

4.3.7.3.3.2 Detailed statistical reasoning and interpretation

Since the effect of Family on the probability of being risk-averse is only sensitive to different time periods from which income parameters are estimated, I only discuss this very sensitivity in detail to avoid unnecessary repetitions. Numerical details on all subgroups are given in Table 47, p. 268.

Detailed comparison of subgroups of dependent variables relating to different ways to gain data input

The relation of Family to the probability of being risk-averse is sensitive to different **time periods from which income parameters are estimated**, namely annual income data of one single year, annual income date from three years around the moving date adjusted for inflation, and annual income data from three years around the moving date not adjusted for inflation (denoted by One, Ad1, and Ad2). First, quasi-complete separation is of no concern in any subgroup since it amounts to only 0.4% of dependent variables at a maximum. Second, applying Leamer's criterion no robust relation for Family is found in any subgroup. Hence, Family is classified as fragile by Leamer's criterion in both subgroups and a statement on significance cannot be made. Third, applying Sala-I-Martin's criterion of robustness, Family is again classified as fragile in both subgroups. Fourth, concerning the tendency for a sign of the coefficient of Family, I observe no tendency for a sign in the subgroup referring to one year's data (One), and a strong tendency for a negative sign in the subgroups referring to three year's data (Ad1 and Ad2).

I interpret these findings as follows: It seems as if the results of Family and its relation to the probability of being risk-averse are sensitive to whether estimations are based on one year's data compared to three year's of data (One versus Ad1 and Ad2), but not sensitive to whether the three year's data are adjusted for inflation or not (Ad1 versus Ad2).

4.3.8 Divorce and risk-attitudes

4.3.8.1 Overview of results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

For about 64% of the 4,536 dependent variables in my study Divorce shows quasi-complete separation. The fact that Divorce almost perfectly predicts risk-attitudes for about 64% of the dependent variables should not be overrated since there are no more than only 12 out of 321 migrants in my sample that are divorced. Therefore, I focus on the statistical analysis of Divorce based on a still sufficiently high sample of 1,645 dependent variables for which no quasi-complete separation occurs.

Based on the sample of 1,645 dependent variables, Divorce does not exhibit a robust influence on the probability of being risk-averse by Leamer's criterion of robustness. Consequently, Divorce is also not significantly related to the probability of being risk-averse in the migration context. Applying Sala-I-Martin's criterion of robustness, Divorce is also classified as fragile, where it is noteworthy that the required 75% of robust relations for Sala-I-Martin's weighted and unweighted criteria are only slightly failed with 73% and 67% robust relations by Sala-I-Martin's weighted and unweighted criterion, respectively. Finally, I derive a strong tendency for a positive sign of Divorce.

Comparison to the literature and interpretation

Since separation as defined by the variable Divorce in my study has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 1,645 dependent variables reveals a strong tendency for a positive sign of Divorce which indicates that Heads that separated from their partner between the wave before and the wave after the move have a higher probability of being risk-averse compared to all other Heads. This is in line with 99.9% of dependent variables for which Divorce perfectly predicts risk-attitudes in the way that all divorced migrants are risk-averse. At the same time the high percentage of dependent variables where Divorce shows quasi-complete separation may indicate that the relation between Divorce and being risk-averse is stronger than suggested by the statistical analysis.

A possible explanation for the higher probability of being risk-averse for divorced migrants might be that the experience of separation and its psychological burden change people's character in the way that they are not willing to take any risk any more.[298]

4.3.8.2 Detailed statistical reasoning

All statistical results on the influence of Divorce on the probability of being risk-averse are summarized in Table 48, p. 273.

[298] Note that further interpretation relating to maintenance obligations cannot be applied in my study for two reasons: First, in my study Divorce relates not only to legal divorce but also to separations of unmarried partnerships. Second, if maintenance obligations were to be paid, divorced Heads could be both the receiving and the paying counterparts. Consequently, the ability to bear risk is of no concern here.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	1,645	64%	0.3%	0%	(-)	73%	87%	(+)	81%	(+)	67%	93%	(+)	76%	(+)
Ind	1,512	61%	0.3%	0%	(-)	72%	86%	(+)	81%	(+)	67%	93%	(+)	76%	(+)
Fam	133	79%	0.0%	n.a.		80%	97%	(+)	77%	(+)	58%	92%	(+)	77%	(+)
Ann	542	64%	0.3%	0%	(-)	72%	87%	(+)	80%	(+)	66%	93%	(+)	76%	(+)
Wor	553	63%	0.3%	0%	(-)	74%	88%	(+)	81%	(+)	67%	93%	(+)	77%	(+)
Lif	550	64%	0.3%	0%	(-)	72%	87%	(+)	81%	(+)	67%	93%	(+)	76%	(+)
Var	1,048	54%	0.0%	n.a.		79%	86%	(+)	80%	(+)	61%	100%	(+)	73%	
LP2	597	74%	0.5%	0%	(-)	61%	91%	(+)	83%	(+)	76%	82%	(+)	83%	(+)
Ed1	386	70%	0.2%	0%	(-)	72%	91%	(+)	81%	(+)	62%	93%	(+)	77%	(+)
Ed2	253	61%	0.9%	0%	(-)	68%	88%	(+)	82%	(+)	63%	93%	(+)	77%	(+)
Ed3	249	62%	0.0%	n.a.		71%	85%	(+)	82%	(+)	70%	93%	(+)	77%	(+)
Ed4	248	62%	0.0%	n.a.		75%	80%	(+)	77%	(+)	71%	93%	(+)	76%	(+)
Ed5	256	60%	0.0%	n.a.		75%	89%	(+)	82%	(+)	68%	93%	(+)	75%	(+)
Ed6	253	61%	0.5%	0%	(-)	75%	89%	(+)	82%	(+)	68%	91%	(+)	75%	(+)
Wei	1,266	44%	0.5%	0%	(-)	76%	85%	(+)	79%	(+)	66%	90%	(+)	73%	
Unw	379	83%	0.0%	n.a.		61%	95%	(+)	88%	(+)	68%	100%	(+)	88%	(+)
Sep	846	63%	0.0%	n.a.		66%	81%	(+)	83%	(+)	51%	100%	(+)	74%	
Poo	799	65%	0.5%	0%	(-)	79%	93%	(+)	79%	(+)	84%	88%	(+)	79%	(+)
One	639	58%	0.0%	n.a.		74%	91%	(+)	82%	(+)	64%	100%	(+)	81%	(+)
Ad1	517	66%	0.4%	0%	(-)	73%	86%	(+)	82%	(+)	70%	89%	(+)	75%	(+)
Ad2	489	68%	0.4%	0%	(-)	70%	84%	(+)	79%	(+)	66%	87%	(+)	71%	
L1	869	43%	0.0%	n.a.		71%	80%	(+)	78%	(+)	61%	99%	(+)	78%	(+)
L2	250	83%	0.0%	n.a.		66%	100%	(+)	93%	(+)	59%	100%	(+)	64%	
Ma	526	65%	0.8%	0%	(-)	79%	93%	(+)	79%	(+)	80%	82%	(+)	79%	(+)
2Kat3	762	50%	0.0%	n.a.		80%	88%	(+)	80%	(+)	78%	93%	(+)	80%	(+)
2Kat1	617	59%	0.4%	0%	(-)	72%	84%	(+)	79%	(+)	65%	93%	(+)	74%	
4Kat2	266	82%	0.4%	0%	(-)	54%	94%	(+)	87%	(+)	38%	91%	(+)	73%	

Table 48: Results of the extreme bounds analysis for the independent variable Divorce aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Sample size and (quasi-) complete separation

Quasi-complete separation is a real issue for the independent variable Divorce. For about 64% of the 4,536 dependent variables (Column (3), Table 48, p. 273) Divorce almost perfectly predicts whether Heads are risk-averse or risk-seeking. The high percentage of dependent variables for which quasi-complete separation occurs is not surprising if it is taken into account that there are no more than 12 out of 321 migrants in my sample that are divorced. This fact also highlights that cases of quasi-complete separation should not be overrated. Therefore, I focus on the statistical analysis of Divorce that is still based on a sufficiently high sample of 1,645 dependent variables (about 36% of 4,536, Column (2), Table 48, p. 273).

Concerning the 2,891 dependent variables (about 64% of 4,536, Column (3), Table 48, p. 273) where quasi-complete separation occurs, it is noteworthy that (i) for 2,888 out of 2,891 all divorced Heads are risk-averse and (ii) for 3 out of the 2,891 only one divorced Head exists and this Head is risk-seeking.

Robustness and significance based on Leamer's criterion

Applying Leamer's criterion of robustness, Divorce is clearly classified as fragile, i.e., Divorce does only show a robust influence on 0.3% of the 1,645 dependent variables (Column (4), Table 48, p. 273). Consequently, Divorce is classified as not significantly related to probability of being risk-averse in the migration context.

Robustness based on Sala-I-Martin's criterion

Applying Sala-I-Martin's criterion of robustness, Divorce slightly fails to be classified as robust, i.e., Sala-I-Martin's weighted (unweighted) criterion finds 73% (67%) of the 1,645 dependent variables to be robustly related to the probability of being risk-averse in the migration context (Columns (7) and (12), Table 48, p. 273).

Tendency for a sign based on Leamer's and Sala-I-Martin's criteria

Irrespective of robustness, a strong tendency for a positive sign of the coefficient of Divorce is observed. This in line with the 99.9% of the dependent variable for which quasi-complete separation occurs where all divorced Heads in the sample are risk-averse.

4.3.8.3 Sensitivity for subgroups of dependent variables

4.3.8.3.1 Overview

Robustness

The fragility of Divorce in explaining the probability of being risk-averse over the 1,645 dependent variables statistically investigated is sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the risk-measure (denoted by Var and LP2), (iii) different education definitions (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iv) different types of clustering (denoted by Sep and Poo), (v) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (vi) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2). Nevertheless, it is noteworthy that the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria are similar in all subgroups where the 75%-criterion is not always met. Therefore, the sensitivity should not be overrated.

However, the fragility of Divorce in explaining the probability of being risk-averse over the 1,645 dependent variables statistically investigated is not sensitive to (i) variations of the planning period (denoted by Ann, Wor, and Lif) and (ii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2).

Tendency for a sign

The finding of a strong tendency for a positive sign of Divorce for the 1,645 dependent variables statistically investigated is not sensitive to any variation investigated in my study. In detail, Divorce is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.8.3.2 Detailed statistical reasoning and interpretation

All subgroups show (i) qualitative similar percentages of dependent variables that exhibit quasi-complete separation, (ii) qualitative similar percentages of dependent variables where Leamer's criterion is met, and (iii) a strong tendency for a positive sign. Although subgroups differ in their statement on robustness by Sala-I-Martin's criterion, the percentages of robust relations are similar in all subgroups. Consequently, the sensitivity of Divorce concerning robustness by Sala-I-Martin's criterion should not be overrated and is therefore not discussed in detail.

4.3.9 On the interpretation of interaction effects in my study

For the interpretation of interaction effects in my study it is important to understand that due to the methodology applied, interaction effects can only be interpreted as partial effects whereas the total effect cannot be observed. The reasons are as follows: First, interaction effects in my study are always entered in the regression together with their main effects. Hence, interaction effects themselves are partial effects. Second, the objective of the extreme bounds analysis is not to estimate the final model, but to decide which variables should be entered in the final model. To achieve this objective, conditional main effects are not needed. Consequently, conditional main effects to estimate total effects are not available from the extreme bounds analysis.

To understand what the non-availability of conditional main effects means for the interpretation of the interaction effects in my study, consider the following example. When the interaction effect MaleDivorce is entered together with its main effects Male and Divorce, the reference category (when all independent variables are null) is women that are not divorced. The coefficients of all main and interaction effects now indicate the effect on the probability of being risk-averse for a certain subgroup of individuals in comparison to the reference category (i.e., women that are not divorced). In detail this means, when all three variables MaleDivorce, Male, and Divorce are entered in the regression, the coefficient of the main effect Male indicates the effect of being male when the person is not divorced (conditional gender-effect), and the coefficient of the main effect Divorce indicates the effect of being divorced when the person is female (conditional divorce-effect). Caveat: The coefficient of the interaction term MaleDivorce does not indicate the total effect of divorced men. Instead, it only indicates a partial effect that can either be interpreted as (i) additional gender-effect for divorced individuals or as (ii) an additional divorce-effect for men. To derive the total effect of divorced men both conditional main effects (Male and Divorce) must be combined with the effect of the interaction term.

Furthermore, a detailed statistical reasoning on the sensitivity of interaction effects in my study will not be given since economic explanations of partial effects without knowing total effects are not possible.

4.3.10 Interactions with Male and their relation to risk-attitudes

In Part C, Section 4.3.2 I found that men tend to possess a higher probability of being risk-averse in the migration context than women, i.e., Male exhibits a strong tendency for a positive sign. In this section I investigate whether this gender-effect depends on other explanatory variables as follows: (i) the number of people moving together and their relation to Head as captured by interaction effects MaleSingle, MalePair, and MaleFamily, (ii) whether Head just separated from his old partner as captured by interaction effect MaleDivorce, (iii) age of Heads as captured by interaction effect MaleAge, and (iv) Head's education level as captured by the interaction effects MaleEdu2 and MaleEdu3.

4.3.10.1 Interaction of Male and single-moves

4.3.10.1.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable MaleSingle captures the interaction effect of being male and being a single-mover. The influence of MaleSingle on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 3,979 dependent variables since quasi-complete separation occurs only for about 12% of the 4,536 dependent variables.[299] Consequently, I focus on the statistical analysis of MaleSingle.

All statistical results on the influence of MaleSingle on the probability of being risk-averse are summarized in Table 49, p. 278. Based on the sample of 3,979 dependent variables, MaleSingle does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, MaleSingle is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a negative sign of MaleSingle.

[299] Concerning the 557 dependent variables (about 12% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 555 out of 557 all male single-movers are risk-averse and (ii) for 2 out of 557 not all but 98% of the male single-movers (85 out of 87) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	3,979	12%	3.2%	0%	(-)	18%	2%	(-)	17%	(-)	16%	0%	(-)	18%	(-)
Ind	3,418	12%	3.3%	0%	(-)	16%	3%	(-)	18%	(-)	15%	0%	(-)	19%	(-)
Fam	561	13%	2.9%	0%	(-)	25%	1%	(-)	8%	(-)	21%	1%	(-)	6%	(-)
Ann	1,325	12%	3.1%	0%	(-)	18%	3%	(-)	17%	(-)	16%	0%	(-)	17%	(-)
Wor	1,327	12%	3.3%	0%	(-)	18%	3%	(-)	17%	(-)	16%	0%	(-)	18%	(-)
Lif	1,327	12%	3.2%	0%	(-)	17%	2%	(-)	17%	(-)	16%	0%	(-)	17%	(-)
Var	2,268	0%	0.1%	0%	(-)	11%	7%	(-)	21%	(-)	9%	0%	(-)	22%	(-)
LP2	1,711	25%	6.3%	0%	(-)	27%	0%	(-)	11%	(-)	25%	0%	(-)	12%	(-)
Ed1	1,119	14%	2.6%	0%	(-)	17%	1%	(-)	16%	(-)	15%	1%	(-)	16%	(-)
Ed2	558	14%	2.8%	0%	(-)	11%	0%	(-)	24%	(-)	9%	0%	(-)	25%	(-)
Ed3	593	8%	6.3%	0%	(-)	31%	0%	(-)	4%	(-)	30%	0%	(-)	5%	(-)
Ed4	593	8%	7.3%	0%	(-)	28%	1%	(-)	9%	(-)	26%	0%	(-)	10%	(-)
Ed5	558	14%	0.0%	n.a.		8%	16%	(-)	25%	(-)	6%	0%	(-)	25%	(-)
Ed6	558	14%	0.9%	0%	(-)	9%	14%	(-)	26%		7%	0%	(-)	26%	
Wei	2,089	8%	2.8%	0%	(-)	13%	1%	(-)	17%	(-)	13%	0%	(-)	19%	(-)
Unw	1,890	17%	3.6%	0%	(-)	22%	3%	(-)	16%	(-)	19%	0%	(-)	16%	(-)
Sep	2,002	12%	3.6%	0%	(-)	17%	0%	(-)	14%	(-)	17%	0%	(-)	15%	(-)
Poo	1,977	13%	2.9%	0%	(-)	18%	4%	(-)	19%	(-)	15%	0%	(-)	20%	(-)
One	1,423	6%	1.5%	0%	(-)	16%	0%	(-)	19%	(-)	15%	0%	(-)	21%	(-)
Ad1	1,297	14%	4.4%	0%	(-)	20%	2%	(-)	14%	(-)	18%	0%	(-)	14%	(-)
Ad2	1,259	17%	3.8%	0%	(-)	16%	5%	(-)	17%	(-)	15%	0%	(-)	17%	(-)
L1	1,512	0%	4.0%	0%	(-)	8%	10%	(-)	25%	(-)	7%	0%	(-)	27%	
L2	1,266	16%	4.0%	0%	(-)	19%	2%	(-)	18%	(-)	17%	0%	(-)	16%	(-)
Ma	1,201	21%	1.7%	0%	(-)	28%	0%	(-)	6%	(-)	25%	0%	(-)	7%	(-)
2Kat3	1,440	5%	1.0%	0%	(-)	10%	0%	(-)	15%	(-)	7%	0%	(-)	16%	(-)
2Kat1	1,327	12%	3.0%	0%	(-)	16%	0%	(-)	13%	(-)	14%	0%	(-)	16%	(-)
4Kat2	1,212	20%	5.6%	0%	(-)	29%	5%	(-)	23%	(-)	28%	0%	(-)	22%	(-)

Table 49: Results of the extreme bounds analysis for the independent variable MaleSingle aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being male and moving alone (as captured by the interaction effect MaleSingle) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 3,979 dependent variables reveals a strong tendency for a negative sign of MaleSingle which can be interpreted in two alternative ways: first, as a strong tendency for an additional negative gender-effect for male single-movers compared to women and male non-single-movers; second, as a strong tendency for an additional negative single-mover effect for men compared women and male non-single-movers. Note that the strong tendency for negative sign of MaleSingle is in line with 99.6% of the remaining 557 dependent variables where (quasi-) complete separation occurs in the way that all male single-movers are risk-averse.

4.3.10.1.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of MaleSingle in explaining the probability of being risk-averse over the 3,979 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, five findings are noteworthy: First, comparing results of subgroups referring to Head's individual versus family income, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to family income (Fam) is about 1.6 times (1.4 times) the percentage of robust relations in the subgroup referring to Head's individual income (Ind). Second, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 2.5 times (2.8 times) the percentage of robust relations in the subgroup referring to variance (Var). Third, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to unweighted samples (Unw) is about 1.7 times (1.5 times) the percentage of robust relations in the subgroup referring to weighted samples (Wei). Fourth, comparing results of subgroups referring to minimizing different predictive errors, I find that when greater weight is put on greater predictive errors step by step (L1, L2, Ma), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criterion increase visibly. Fifth, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criterion increase visibly.

Tendency for a sign

The finding of a strong tendency for a negative sign of MaleSingle for the 3,979 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, MaleSingle is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.10.2 Interaction of Male and pair-moves

4.3.10.2.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable MalePair captures the interaction effect of being male and being a pair-mover. The influence of MalePair on the probability of being risk-averse can be statistically investigated for 4,387 dependent variables since quasi-complete separation only occurs for about 3% of the 4,536 dependent variables.[300] Consequently, I focus on the statistical analysis of MalePair.

All statistical results on the influence of MalePair on the probability of being risk-averse are summarized in Table 50, p. 281. Based on the sample of 4,387 dependent variables, MalePair does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, MalePair is also not significantly related to the probability of being risk-averse. In addition, I observe no tendency for sign of MalePair.

[300] Concerning the 149 dependent variables (about 3% of 4,536) where quasi-complete separation occurs, it is noteworthy that all male pair-movers in the sample are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,387	3%	0.6%	0%	(-)	28%	70%		47%		16%	58%		49%	
Ind	3,742	4%	0.6%	0%	(-)	28%	73%		49%		16%	63%		51%	
Fam	645	0%	0.2%	0%	(-)	26%	51%		36%		16%	29%		38%	
Ann	1,465	3%	0.5%	0%	(-)	28%	71%		46%		17%	59%		49%	
Wor	1,461	3%	0.6%	0%	(-)	28%	69%		47%		16%	57%		50%	
Lif	1,461	3%	0.5%	0%	(-)	28%	71%		48%		16%	59%		49%	
Var	2,233	2%	0.0%	n.a.		25%	82%	(+)	48%		15%	97%	(+)	51%	
LP2	2,154	5%	1.1%	0%	(-)	31%	60%		46%		17%	22%	(-)	47%	
Ed1	1,270	2%	0.8%	0%	(-)	27%	61%		41%		17%	40%		44%	
Ed2	625	4%	1.9%	0%	(-)	29%	68%		46%		18%	51%		50%	
Ed3	621	4%	0.0%	n.a.		25%	77%	(+)	53%		16%	75%	(+)	53%	
Ed4	621	4%	0.0%	n.a.		29%	70%		52%		17%	74%		52%	
Ed5	625	4%	0.0%	n.a.		28%	77%	(+)	48%		15%	65%		50%	
Ed6	625	4%	0.5%	0%	(-)	30%	77%	(+)	48%		16%	66%		50%	
Wei	2,154	5%	0.5%	0%	(-)	35%	80%	(+)	57%		18%	80%	(+)	59%	
Unw	2,233	2%	0.6%	0%	(-)	21%	54%		38%		14%	31%		40%	
Sep	2,220	2%	0.6%	0%	(-)	28%	63%		43%		19%	44%		42%	
Poo	2,167	4%	0.5%	0%	(-)	28%	77%	(+)	51%		14%	78%	(+)	56%	
One	1,512	0%	0.1%	0%	(-)	34%	74%		48%		7%	80%	(+)	52%	
Ad1	1,441	5%	0.6%	0%	(-)	26%	69%		48%		22%	57%		48%	
Ad2	1,434	5%	1.0%	0%	(-)	23%	65%		45%		20%	52%		47%	
L1	1,512	0%	0.0%	n.a.		25%	95%	(+)	55%		17%	85%	(+)	60%	
L2	1,512	0%	0.0%	n.a.		32%	59%		40%		11%	41%		38%	
Ma	1,363	10%	1.7%	0%	(-)	27%	58%		46%		21%	44%		50%	
2Kat3	1,512	0%	0.8%	0%	(-)	17%	68%		34%		12%	54%		38%	
2Kat1	1,470	3%	0.0%	n.a.		40%	79%	(+)	54%		20%	72%		55%	
4Kat2	1,405	7%	0.9%	0%	(-)	28%	58%		55%		17%	44%		55%	

Table 50: Results of the extreme bounds analysis for the independent variable MalePair aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being male and moving with one's partner (as captured by the interaction effect MalePair) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,387 dependent variables reveals a fragile effect of MalePair on the probability of being risk-averse with no tendency for a sign. This can be interpreted in two alternative ways: first, as no additional gender-effect for male pair-movers compared to women and male non-pair-movers; second, as no additional single-mover effect for men compared to women and male non-pair-movers.

4.3.10.2.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of MalePair in explaining the probability of being risk-averse over the 4,387 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, one finding is noteworthy: Comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to weighted samples (Wei) is about 1.7 times (1.3 times) the percentage of robust relations in the subgroup referring to unweighted samples (Unw).

Tendency for a sign

The finding of no tendency for a sign of MalePair for the 4,387 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the risk-measure (denoted by Var and LP2), and not sensitive to (iv) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2).

However, the finding of no tendency for a sign of MalePair is sensitive to (i) variations of the education definition (denoted by Ed1 to Ed6), (ii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iii) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and

Poo), (iv) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and sensitive to (v) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.10.3 Interaction of Male and family-moves

4.3.10.3.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable MaleFamily captures the interaction effect of being male and being a family-mover. The influence of MaleFamily on the probability of being risk-averse can be statistically investigated for 4,303 dependent variables since quasi-complete separation only occurs for about 5% of the 4,536 dependent variables.[301] Consequently, I focus on the statistical analysis of MaleFamily.

All statistical results on the influence of MaleFamily on the probability of being risk-averse are summarized in Table 51, p. 284. Based on the sample of 4,303 dependent variables, MaleFamily does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, MaleFamily is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a positive sign of MaleFamily.

[301] Concerning the 233 dependent variables (about 0.1% of 4,536) where quasi-complete separation occurs, it is noteworthy that all male family-movers in the sample are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,303	5%	1.3%	100%	(+)	25%	87%	(+)	84%	(+)	24%	86%	(+)	85%	(+)
Ind	3,693	5%	1.3%	100%	(+)	23%	85%	(+)	82%	(+)	23%	83%	(+)	84%	(+)
Fam	610	6%	1.4%	100%	(+)	32%	97%	(+)	92%	(+)	30%	99%	(+)	94%	(+)
Ann	1,433	5%	1.3%	100%	(+)	25%	87%	(+)	84%	(+)	23%	85%	(+)	85%	(+)
Wor	1,435	5%	1.3%	100%	(+)	25%	87%	(+)	84%	(+)	24%	86%	(+)	86%	(+)
Lif	1,435	5%	1.3%	100%	(+)	24%	87%	(+)	83%	(+)	23%	85%	(+)	85%	(+)
Var	2,236	1%	0.0%	n.a.		10%	94%	(+)	78%	(+)	9%	100%	(+)	81%	(+)
LP2	2,067	9%	2.6%	100%	(+)	41%	85%	(+)	90%	(+)	40%	82%	(+)	89%	(+)
Ed1	1,225	5%	1.2%	100%	(+)	26%	90%	(+)	85%	(+)	24%	91%	(+)	87%	(+)
Ed2	615	5%	1.4%	100%	(+)	21%	79%	(+)	76%	(+)	19%	80%	(+)	79%	(+)
Ed3	618	5%	2.3%	100%	(+)	41%	96%	(+)	95%	(+)	41%	94%	(+)	96%	(+)
Ed4	618	5%	2.0%	100%	(+)	30%	94%	(+)	91%	(+)	32%	92%	(+)	90%	(+)
Ed5	615	5%	0.5%	100%	(+)	14%	68%		77%	(+)	13%	54%		78%	(+)
Ed6	612	6%	0.8%	100%	(+)	15%	67%		76%	(+)	13%	60%		78%	(+)
Wei	2,151	5%	0.5%	100%	(+)	18%	99%	(+)	86%	(+)	16%	100%	(+)	87%	(+)
Unw	2,152	5%	2.1%	100%	(+)	31%	80%	(+)	82%	(+)	31%	78%	(+)	84%	(+)
Sep	2,169	4%	1.7%	100%	(+)	25%	96%	(+)	86%	(+)	27%	96%	(+)	89%	(+)
Poo	2,134	6%	0.9%	100%	(+)	24%	78%	(+)	81%	(+)	21%	72%		81%	(+)
One	1,429	5%	1.3%	100%	(+)	25%	87%	(+)	81%	(+)	22%	86%	(+)	81%	(+)
Ad1	1,440	5%	1.3%	100%	(+)	25%	89%	(+)	86%	(+)	24%	89%	(+)	89%	(+)
Ad2	1,434	5%	1.5%	100%	(+)	23%	86%	(+)	84%	(+)	25%	81%	(+)	86%	(+)
L1	1,512	0%	2.6%	100%	(+)	11%	93%	(+)	79%	(+)	9%	100%	(+)	79%	(+)
L2	1,501	1%	1.3%	100%	(+)	33%	85%	(+)	85%	(+)	33%	86%	(+)	89%	(+)
Ma	1,290	15%	0.1%	100%	(+)	30%	88%	(+)	88%	(+)	30%	80%	(+)	89%	(+)
2Kat3	1,473	3%	0.0%	n.a.		16%	95%	(+)	89%	(+)	15%	94%	(+)	90%	(+)
2Kat1	1,449	4%	0.1%	100%	(+)	24%	83%	(+)	85%	(+)	23%	78%	(+)	85%	(+)
4Kat2	1,381	9%	3.9%	100%	(+)	35%	86%	(+)	76%	(+)	34%	86%	(+)	80%	(+)

Table 51: Results of the extreme bounds analysis for the independent variable MaleFamily aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being male and moving with one's family (as captured by the interaction effect MaleFamily) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,303 dependent variables reveals a fragile effect of MaleFamily on the probability of being risk-averse with a strong tendency for a positive sign. This can be interpreted in two alternative ways: first, as a strong tendency for an additional positive gender-effect for male family-movers compared to women and male non-family-movers; second, as a strong tendency for an additional positive family-mover effect for men compared to women and male non-family-movers.

4.3.10.3.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of MaleFamily in explaining the probability of being risk-averse over the 4,303 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three findings are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 4.1 times (4.4 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing result of subgroups referring to weighted/unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to unweighted sample (Unw) is about 1.7 times (1.9 times) the percentage of robust relations in the subgroup referring to weighted samples (Wei). Third, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a strong tendency for a positive sign of MaleFamily for the 4,303 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, MaleFamily is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the risk-measure (denoted by Var and LP2), (iv) variations of the education definition (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2), (viii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (iv) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.10.4 Interaction of Male and Divorce

4.3.10.4.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable MaleDivorce captures the interaction effect of being male and being divorced. For about 67% of the 4,536 dependent variables in my study MaleDivorce shows quasi-complete separation.[302] The fact that MaleDivorce almost perfectly predicts risk-attitudes for about 67% of the dependent variables should not be overrated since there are no more than only 6 out of 321 migrants in my sample that are divorced men. Therefore, I focus on the statistical analysis of MaleDivorce based on a still sufficiently high sample of 1,492 dependent variables for which no quasi-complete separation occurs.

All statistical results on the influence of MaleDivorce on the probability of being risk-averse are summarized in Table 52, p. 287. Based on 1,492 dependent variables, MaleDivorce does not exhibit a robust influence on the probability of being risk-averse by Leamer's criterion. Consequently, MaleDivorce is also not significantly related to the probability of being risk-averse. In contrast, applying Sala-I-Martin's criterion of robustness, MaleDivorce is clearly classified as robust. Finally, I derive a strong tendency for a negative sign of MaleDivorce.

[302] Concerning the 3,044 dependent variables (about 67% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 2,973 out of 3,044 all divorced men are risk-averse, (ii) for 50 out of 3,044 there is no divorced man in the sample, and (iii) for 21 out of 3,044 only up to two divorced men exist in the sample and these men are risk-seeking.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	1,492	67%	0.9%	0%	(-)	99%	11%	(-)	11%	(-)	100%	11%	(-)	11%	(-)
Ind	1,386	64%	0.8%	0%	(-)	99%	12%	(-)	12%	(-)	99%	12%	(-)	12%	(-)
Fam	106	84%	0.9%	0%	(-)	99%	0%	(-)	0%	(-)	100%	0%	(-)	0%	(-)
Ann	491	68%	0.9%	0%	(-)	99%	12%	(-)	11%	(-)	99%	11%	(-)	11%	(-)
Wor	502	67%	0.9%	0%	(-)	100%	11%	(-)	11%	(-)	100%	11%	(-)	11%	(-)
Lif	499	67%	0.9%	0%	(-)	99%	11%	(-)	11%	(-)	99%	11%	(-)	11%	(-)
Var	994	56%	0.7%	0%	(-)	99%	17%	(-)	17%	(-)	99%	16%	(-)	17%	(-)
LP2	498	78%	1.1%	0%	(-)	100%	0%	(-)	0%	(-)	100%	0%	(-)	0%	(-)
Ed1	338	74%	0.7%	0%	(-)	99%	8%	(-)	8%	(-)	99%	8%	(-)	8%	(-)
Ed2	232	64%	1.9%	0%	(-)	100%	12%	(-)	12%	(-)	100%	12%	(-)	12%	(-)
Ed3	231	64%	0.9%	0%	(-)	99%	13%	(-)	13%	(-)	99%	12%	(-)	13%	(-)
Ed4	230	65%	0.9%	0%	(-)	99%	13%	(-)	13%	(-)	99%	12%	(-)	13%	(-)
Ed5	232	64%	0.5%	0%	(-)	99%	12%	(-)	12%	(-)	100%	12%	(-)	12%	(-)
Ed6	229	65%	0.5%	0%	(-)	100%	12%	(-)	12%	(-)	100%	12%	(-)	12%	(-)
Wei	1,149	49%	1.6%	0%	(-)	99%	11%	(-)	11%	(-)	99%	11%	(-)	11%	(-)
Unw	343	85%	0.1%	0%	(-)	100%	10%	(-)	10%	(-)	100%	10%	(-)	10%	(-)
Sep	807	64%	1.5%	0%	(-)	100%	14%	(-)	14%	(-)	100%	13%	(-)	14%	(-)
Poo	685	70%	0.3%	0%	(-)	99%	8%	(-)	8%	(-)	100%	8%	(-)	8%	(-)
One	585	61%	1.6%	0%	(-)	98%	16%	(-)	15%	(-)	99%	15%	(-)	15%	(-)
Ad1	463	69%	0.5%	0%	(-)	100%	8%	(-)	8%	(-)	100%	8%	(-)	8%	(-)
Ad2	444	71%	0.5%	0%	(-)	100%	9%	(-)	9%	(-)	99%	8%	(-)	9%	(-)
L1	824	46%	0.1%	0%	(-)	100%	20%	(-)	20%	(-)	100%	20%	(-)	20%	(-)
L2	250	83%	0.7%	0%	(-)	100%	0%	(-)	0%	(-)	100%	0%	(-)	0%	(-)
Ma	418	72%	1.8%	0%	(-)	98%	0%	(-)	0%	(-)	99%	0%	(-)	0%	(-)
2Kat3	699	54%	0.4%	0%	(-)	100%	13%	(-)	13%	(-)	100%	13%	(-)	13%	(-)
2Kat1	557	63%	0.9%	0%	(-)	100%	13%	(-)	13%	(-)	100%	13%	(-)	13%	(-)
4Kat2	236	84%	1.3%	0%	(-)	96%	2%	(-)	2%	(-)	97%	0%	(-)	2%	(-)

Table 52: Results of the extreme bounds analysis for the independent variable MaleDivorce aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being male and being divorced (as captured by the interaction effect MaleDivorce) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 1,492 dependent variables reveals a strong tendency for a negative sign of MaleDivorce that can be interpreted in two alternative ways: first, as a strong tendency for an additional negative gender-effect for divorced men compared to women and non-divorced men; second, as a strong tendency for an additional negative divorce-effect for men compared to women and non-divorced men.

Note that (i) the robust influence of MaleDivorce on the probability of being risk-averse by Sala-I-Martin's criterion and (ii) the strong tendency for a negative sign of MaleDivorce derived from the statistical analysis of 1,492 dependent variables are not in line with 97.7% of the remaining 3,044 dependent variables where quasi-complete separation occurs in the way that all divorced men are risk-averse. Therefore, findings of the statistical analysis must be interpreted with caution.

4.3.10.4.2 Sensitivity for subgroups of dependent variables

Robustness

First, the fragility and non-significance of MaleDivorce by Leamer's criterion and second, the robustness of MaleDivorce by Sala-I-Martin's criterion in explaining the probability of being risk-averse over the 1,492 dependent variables statistically investigated are not sensitive to different ways to estimate risk-attitudes.

Tendency for a sign

The finding of a strong tendency for a negative sign of MaleDivorce for the 1,492 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, MaleDivorce is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the risk-measure (denoted by Var and LP2), (iv) variations of the education definition (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2), (viii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (iv) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.10.5 Interaction of Male and Age

4.3.10.5.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable MaleAge captures the interaction effect of gender and age. The influence of MaleAge on the probability of being risk-averse can be statistically investigated for 4,478 dependent variables since quasi-complete separation only occurs for about 1% of the 4,536 dependent variables.[303] Consequently, I focus on the statistical analysis of MaleAge.

All statistical results on the influence of MaleAge on the probability of being risk-averse are summarized in Table 53, p. 290. Based on the sample of 4,478 dependent variables, MaleAge does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, MaleAge is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a positive sign of MaleAge.

[303] Concerning the 58 dependent variables (about 1% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 48 out of 58 all men are risk-averse and (ii) for 10 out of 58 all men with ages other than 30, 47 or 50 are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,478	1%	0.4%	100%	(+)	18%	98%	(+)	74%		17%	99%	(+)	74%	
Ind	3,840	1%	0.3%	100%	(+)	18%	98%	(+)	74%		17%	99%	(+)	74%	
Fam	638	2%	0.5%	100%	(+)	18%	97%	(+)	71%		20%	99%	(+)	74%	
Ann	1,492	1%	0.5%	100%	(+)	19%	99%	(+)	74%		18%	99%	(+)	74%	
Wor	1,494	1%	0.3%	100%	(+)	18%	98%	(+)	74%		17%	100%	(+)	74%	
Lif	1,492	1%	0.3%	100%	(+)	19%	98%	(+)	74%		18%	99%	(+)	74%	
Var	2,268	0%	0.3%	100%	(+)	11%	99%	(+)	63%		12%	100%	(+)	65%	
LP2	2,210	3%	0.4%	100%	(+)	26%	98%	(+)	85%	(+)	23%	99%	(+)	83%	(+)
Ed1	1,274	2%	0.2%	100%	(+)	16%	99%	(+)	70%		16%	100%	(+)	72%	
Ed2	636	2%	0.0%	n.a.		15%	100%	(+)	71%		14%	100%	(+)	71%	
Ed3	648	0%	0.6%	100%	(+)	27%	97%	(+)	85%	(+)	26%	99%	(+)	85%	(+)
Ed4	648	0%	1.4%	100%	(+)	31%	97%	(+)	85%	(+)	28%	99%	(+)	85%	(+)
Ed5	636	2%	0.0%	n.a.		11%	100%	(+)	67%		11%	100%	(+)	68%	
Ed6	636	2%	0.0%	n.a.		12%	100%	(+)	68%		11%	100%	(+)	67%	
Wei	2,220	2%	0.4%	100%	(+)	14%	100%	(+)	68%		11%	100%	(+)	69%	
Unw	2,258	0%	0.3%	100%	(+)	23%	97%	(+)	80%	(+)	23%	99%	(+)	79%	(+)
Sep	2,258	0%	0.3%	100%	(+)	24%	98%	(+)	79%	(+)	22%	99%	(+)	80%	(+)
Poo	2,220	2%	0.4%	100%	(+)	13%	99%	(+)	69%		12%	100%	(+)	68%	
One	1,512	0%	0.5%	100%	(+)	18%	99%	(+)	85%	(+)	15%	100%	(+)	84%	(+)
Ad1	1,483	2%	0.4%	100%	(+)	19%	98%	(+)	68%		18%	100%	(+)	68%	
Ad2	1,483	2%	0.2%	100%	(+)	17%	98%	(+)	69%		18%	99%	(+)	70%	
L1	1,512	0%	0.2%	100%	(+)	6%	99%	(+)	54%		6%	100%	(+)	57%	
L2	1,512	0%	0.1%	100%	(+)	23%	100%	(+)	91%	(+)	24%	100%	(+)	93%	(+)
Ma	1,454	4%	0.7%	100%	(+)	26%	96%	(+)	78%	(+)	22%	98%	(+)	72%	
2Kat3	1,512	0%	0.0%	n.a.		8%	100%	(+)	66%		9%	100%	(+)	68%	
2Kat1	1,484	2%	0.5%	100%	(+)	13%	99%	(+)	73%		14%	100%	(+)	74%	
4Kat2	1,482	2%	0.5%	100%	(+)	34%	97%	(+)	83%	(+)	29%	99%	(+)	81%	(+)

Table 53: Results of the extreme bounds analysis for the independent variable MaleAge aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of gender and age (as captured by the interaction effect MaleAge) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,478 dependent variables reveals a strong tendency for a positive sign of MaleAge that can be interpreted in two alternative ways: first, as a strong tendency for an additional positive gender-effect for men due to an additional year of age compared to women; second, as a strong tendency for an additional positive age-effect for men compared to women.

4.3.10.5.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of MaleAge in explaining the probability of being risk-averse over the 4,478 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, four findings are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 2.4 times (1.9 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing results of subgroups referring to weighted/unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to unweighted samples (Unw) is about 1.6 times (2.1 times) the percentage of robust relations in the subgroup referring to weighted samples (Wei). Third, comparing results of subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to separate clustering in each year (Sep) is about 1.8 times (1.8 times) the percentage of robust relations in the subgroup referring to pooled clustering (Poo). Fourth, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a strong tendency for a positive sign of MaleAge for the 4,478 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, MaleAge is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) variations of the risk-measure (denoted by Var and LP2), (iv) variations of the education definition (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2), (viii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (iv) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.10.6 Interaction of Male and Edu2

4.3.10.6.1 Results irrespective of the way risk-attitudes are estimated

<u>Significance, robustness, and tendency for a sign</u>

The dummy variable MaleEdu2 captures the interaction effect of gender and having a high school diploma or an associate's degree as highest education level. The influence of MaleEdu2 on the probability of being risk-averse can be statistically investigated for 4,433 dependent variables since quasi-complete separation only occurs for about 2% of the 4,536 dependent variables.[304] Consequently, I focus on the statistical analysis of MaleEdu2.

All statistical results on the influence of MaleEdu2 on the probability of being risk-averse are summarized in Table 54, p. 293. Based on the sample of 4,433 dependent variables, MaleEdu2 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Consequently, MaleEdu2 is also not significantly related to the probability of being risk-averse in the migration context. Yet, I derive a vague tendency for a negative sign of MaleEdu2.

[304] Concerning the 103 dependent variables (about 2% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 48 out of 103 all men with education level two (MaleEdu2=1) and almost all other migrants (MaleEdu2=0) are risk-averse and (ii) for 55 out of 103 only three risk-seeking migrants exist.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	**number of dependent variables**	**missing values**	**Leamer**			**Sala-I-Martin weighted**					**Sala-I-Martin unweighted**				
			significant	**positive**		**CDF(0) ≥95%**	**positive**		**positive**		**CDF(0) ≥95%**	**positive**		**positive**	
	n	**% of 4,536**	**% of n**	**% of significant**		**% of n**	**% of robust**		**% of n**		**% of n**	**% of robust**		**% of n**	
Total	4,433	2%	2.4%	0%	(-)	36%	49%		37%		36%	41%		34%	
Ind	3,816	2%	2.1%	0%	(-)	36%	50%		37%		36%	41%		34%	
Fam	617	5%	4.8%	0%	(-)	36%	47%		37%		39%	43%		37%	
Ann	1,477	2%	2.2%	0%	(-)	36%	50%		36%		37%	40%		34%	
Wor	1,479	2%	2.8%	0%	(-)	37%	50%		37%		36%	42%		35%	
Lif	1,477	2%	2.3%	0%	(-)	36%	49%		37%		36%	42%		35%	
Var	2,268	0%	1.2%	0%	(-)	24%	46%		36%		22%	33%		34%	
LP2	2,165	5%	3.7%	0%	(-)	50%	51%		38%		51%	45%		35%	
Ed1	1,253	3%	3.9%	0%	(-)	37%	52%		42%		37%	49%		41%	
Ed2	636	2%	1.9%	0%	(-)	36%	53%		38%		32%	50%		36%	
Ed3	636	2%	2.5%	0%	(-)	43%	48%		31%		42%	39%		29%	
Ed4	636	2%	1.4%	0%	(-)	42%	57%		38%		42%	45%		34%	
Ed5	636	2%	1.9%	0%	(-)	29%	43%		37%		35%	30%		36%	
Ed6	636	2%	1.9%	0%	(-)	30%	35%		28%		28%	22%	(-)	25%	(-)
Wei	2,187	4%	1.1%	0%	(-)	40%	51%		33%		35%	47%		33%	
Unw	2,246	1%	3.8%	0%	(-)	32%	47%		40%		37%	36%		36%	
Sep	2,246	1%	1.3%	0%	(-)	35%	58%		42%		36%	53%		41%	
Poo	2,187	4%	3.6%	0%	(-)	37%	41%		31%		36%	29%		28%	
One	1,512	0%	5.6%	0%	(-)	34%	45%		29%		30%	16%	(-)	23%	(-)
Ad1	1,462	3%	0.3%	0%	(-)	34%	57%		41%		35%	56%		41%	
Ad2	1,459	4%	1.4%	0%	(-)	41%	47%		40%		44%	47%		40%	
L1	1,509	0%	1.2%	0%	(-)	38%	45%		45%		37%	53%		44%	
L2	1,506	0%	1.3%	0%	(-)	37%	53%		32%		38%	48%		34%	
Ma	1,418	6%	4.9%	0%	(-)	33%	51%		33%		34%	19%	(-)	25%	(-)
2Kat3	1,512	0%	0.9%	0%	(-)	28%	32%		30%		29%	19%	(-)	27%	
2Kat1	1,478	2%	1.3%	0%	(-)	38%	62%		40%		40%	49%		35%	
4Kat2	1,443	5%	5.2%	0%	(-)	42%	50%		41%		41%	49%		42%	

Table 54: Results of the extreme bounds analysis for the independent variable MaleEdu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being male and having a high school diploma or an associate's degree as highest education level (as captured by the interaction effect MaleEdu2) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,433 dependent variables reveals a vague tendency for a negative sign of MaleEdu2 that can be interpreted in two alternative ways: first, as a vague tendency for an additional negative gender-effect for men having a high school diploma or an associate's degree (Edu2=1) compared to women and men having other education levels (i.e., men having less than a high school diploma (Edu1=1) and men having a bachelor's degree or higher (Edu3=1)); second, as a vague tendency for an additional negative education-effect of having a high school diploma or an associate's degree (Edu2=1) for men compared to women and men having other education levels (i.e., ,men having less than a high school diploma (Edu1=1) and men having a bachelor's degree or higher (Edu3=1)).

4.3.10.6.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of MaleEdu2 in explaining the probability of being risk-averse over the 4,433 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 2.1 times (2.3 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a vague tendency for a negative sign of MaleEdu2 for the 4,433 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), and not sensitive to (iii) variations in the measurement of predictive errors (denoted by L1, L2, and Ma).

However, the finding of a vague tendency for a negative sign of MaleEdu2 is sensitive to (i) variations of the risk-measure (denoted by Var and LP2), (ii) variations of the education definition (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iv) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (v) whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2), and sensitive to (vi) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.10.7 Interaction of Male and Edu3

4.3.10.7.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable MaleEdu3 captures the interaction effect of gender and having a bachelor's degree or higher as highest education level. The influence of MaleEdu3 on the probability of being risk-averse can be statistically investigated for 4,148 dependent variables since quasi-complete separation occurs for about 9% of the 4,536 dependent variables.[305] Consequently, I focus on the statistical analysis of MaleEdu3.

All statistical results on the influence of MaleEdu3 on the probability of being risk-averse are summarized in Table 55, p. 296. Based on the sample of 4,148 dependent variables, MaleEdu3 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Consequently, MaleEdu3 is also not significantly related to the probability of being risk-averse in the migration context. Yet, I derive a strong tendency for a negative sign of MaleEdu3.

[305] Concerning the 388 dependent variables (about 9% of 4,536) where quasi-complete separation occurs, it is noteworthy that all men with education level three (MaleEdu3=1) and almost all other migrants (MaleEdu3=0) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,148	9%	15.7%	1%	(-)	32%	30%		50%		39%	24%	(-)	48%	
Ind	3,580	8%	13.9%	1%	(-)	31%	32%		50%		37%	25%	(-)	48%	
Fam	568	12%	26.4%	1%	(-)	38%	21%	(-)	47%		49%	17%	(-)	46%	
Ann	1,384	8%	15.5%	0%	(-)	32%	31%		50%		39%	24%	(-)	48%	
Wor	1,379	9%	15.9%	1%	(-)	32%	30%		49%		40%	24%	(-)	48%	
Lif	1,385	8%	15.6%	0%	(-)	31%	29%		50%		38%	23%	(-)	48%	
Var	2,146	5%	0.3%	57%		17%	63%		62%		21%	52%		60%	
LP2	2,002	12%	31.0%	0%	(-)	49%	18%	(-)	37%		58%	13%	(-)	35%	
Ed1	1,175	9%	22.7%	0%	(-)	36%	21%	(-)	47%		45%	18%	(-)	45%	
Ed2	607	6%	18.8%	0%	(-)	31%	20%	(-)	44%		40%	19%	(-)	43%	
Ed3	593	8%	18.5%	0%	(-)	34%	29%		48%		41%	21%	(-)	45%	
Ed4	593	8%	18.5%	0%	(-)	37%	32%		52%		43%	25%	(-)	51%	
Ed5	590	9%	4.0%	0%	(-)	26%	49%		55%		31%	36%		54%	
Ed6	590	9%	4.5%	14%	(-)	25%	46%		55%		29%	35%		53%	
Wei	2,053	9%	16.2%	1%	(-)	34%	44%		60%		40%	39%		58%	
Unw	2,095	8%	15.1%	0%	(-)	30%	14%	(-)	39%		38%	8%	(-)	38%	
Sep	2,146	5%	17.9%	1%	(-)	32%	19%	(-)	48%		36%	16%	(-)	47%	
Poo	2,002	12%	13.5%	0%	(-)	32%	42%		51%		43%	31%		49%	
One	1,322	13%	0.5%	63%		34%	76%	(+)	78%	(+)	36%	71%		78%	(+)
Ad1	1,425	6%	23.9%	0%	(-)	32%	5%	(-)	37%		41%	4%	(-)	34%	
Ad2	1,401	7%	22.6%	0%	(-)	30%	7%	(-)	36%		40%	4%	(-)	33%	
L1	1,441	5%	17.8%	0%	(-)	28%	21%	(-)	41%		34%	16%	(-)	41%	
L2	1,429	5%	21.2%	0%	(-)	31%	30%		59%		37%	21%	(-)	55%	
Ma	1,278	15%	8.0%	4%	(-)	37%	37%		50%		47%	32%		48%	
2Kat3	1,512	0%	12.6%	2%	(-)	26%	26%		45%		39%	23%	(-)	43%	
2Kat1	1,445	4%	15.6%	0%	(-)	35%	39%		55%		39%	32%		52%	
4Kat2	1,191	21%	18.8%	0%	(-)	36%	23%	(-)	49%		38%	14%	(-)	49%	

Table 55: Results of the extreme bounds analysis for the independent variable MaleEdu3 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being male and having a bachelor's degree or higher as highest education level (as captured by the interaction effect MaleEdu3) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,148 dependent variables reveals a strong tendency for a negative sign of MaleEdu3 that can be interpreted in two alternative ways: first, as a strong tendency for an additional negative gender-effect for men having a bachelor's degree or higher (Edu3=1) compared to women and men having other education levels (i.e., men having less than a high school diploma (Edu1=1) and men having a high school diploma or an associate's degree (Edu2=1)); second, as a strong tendency for an additional negative education-effect of having a bachelor's degree or higher (Edu3=1) for men compared to women and men having other education levels (i.e., men having less than a high school diploma (Edu1=1) and men having a high school diploma or an associate's degree (Edu2=1)).

4.3.10.7.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of MaleEdu3 in explaining the probability of being risk-averse over the 4,148 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three findings are noteworthy: First, comparing the results of the subgroups referring to Head's versus family income, I find that the percentage of robust relations by Leamer's/Sala-I-Martin's weighted/ Sala-I-Martin's unweighted criterion in the subgroup referring to family income (Fam) is about 1.9 times/1.2 times/1.3 times the percentage of robust relations in the subgroup referring to Head's individual income (Ind). Second, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Leamer's/Sala-I-Martin's weighted/ Sala-I-Martin's unweighted criterion in the subgroup referring to semi-variance (LP2) is about 103 times/2.9 times/2.8 times the percentage of robust relations in the subgroup referring to variance (Var). Third, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-attitude are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by all criterion of robustness increase visibly.

Tendency for a sign

The finding of a strong tendency for a negative sign of MaleEdu3 for the 4,148 dependent variables statistically investigated is not sensitive to (i) whether the migration decision is based on Head's individual income or family income (denoted by Ind and Fam), and (ii) variations of the planning period (denoted by Ann, Wor, and Lif).

However, the finding of a strong tendency for a negative sign of MaleEdu3 is sensitive to (i) variations of the risk-measure (denoted by Var and LP2), (ii) variations of the education definition (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iv) which type of clustering is performed to cluster people from which income parameters are estimated (denoted by Sep and Poo), (v) whether income parameters are estimated from annual income data of one or three years of data (denoted by One, Ad1 and Ad2), (vi) variations in the measurement of predictive errors (denoted by L1, L2, and Ma), and sensitive to (vii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11 Interactions with Single, Pair, Family and their relation to risk-attitudes

In Part C, Section 4.3.7 I found that single-movers tend to possess a higher probability of being risk-averse compared to all other migrants (i.e., Single exhibits a strong tendency for a positive sign), family-movers tend to possess a lower probability of being risk-averse than all other migrants (i.e., Family exhibits a strong tendency for a negative sign), while pair-movers tend to be not systematically different from all other migrants (i.e., Pair exhibits no tendency for a sign). In this section, I investigate whether these effects (i.e., single-, pair-, and family-effects) depend on other explanatory variables, namely (i) Head's age as captured by interaction effects with Age and (ii) Head's education level as captured by interaction effects with Edu2 and Edu3, respectively.

4.3.11.1 Interaction of single-moves and Age

4.3.11.1.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable SingleAge captures the interaction effect of Head being a single-mover and Head's age. The influence of SingleAge on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,272 dependent variables since (quasi-) complete separation occurs only for about 6% of the 4,536 dependent variables.[306] Consequently, I focus on the statistical analysis of SingleAge.

All statistical results on the influence of SingleAge on the probability of being risk-averse are summarized in Table 56, p. 300. Based on the sample of 4,272 dependent variables, SingleAge does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, SingleAge is also not significantly related to the probability of being risk-averse. Yet, I observe a vague tendency for a negative sign of SingleAge.

[306] Concerning the 264 dependent variables (about 6% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 260 out of 264 all single-movers (SingleAge=0) are risk-averse and (ii) for 4 out of 264 not all but 97% to 99% of the single-movers (2 to 4 out of 138) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,272	6%	0.9%	0%	(-)	22%	39%		43%		20%	34%		37%	
Ind	3,698	5%	1.0%	0%	(-)	21%	40%		42%		18%	36%		37%	
Fam	574	11%	0.2%	0%	(-)	30%	30%		44%		28%	27%		41%	
Ann	1,421	6%	0.9%	0%	(-)	21%	37%		43%		18%	34%		37%	
Wor	1,426	6%	0.9%	0%	(-)	23%	38%		42%		21%	33%		37%	
Lif	1,425	6%	0.8%	0%	(-)	23%	40%		43%		20%	36%		37%	
Var	2,268	0%	1.7%	0%	(-)	26%	17%	(-)	30%		25%	15%	(-)	28%	
LP2	2,004	12%	0.0%	n.a.		18%	75%	(+)	57%		13%	74%		47%	
Ed1	1,183	9%	0.8%	0%	(-)	27%	31%		42%		24%	26%		38%	
Ed2	609	6%	0.9%	0%	(-)	22%	36%		41%		19%	35%		35%	
Ed3	611	6%	0.9%	0%	(-)	18%	49%		45%		15%	44%		40%	
Ed4	611	6%	0.8%	0%	(-)	16%	58%		50%		14%	49%		43%	
Ed5	629	3%	0.9%	0%	(-)	23%	37%		38%		22%	32%		34%	
Ed6	629	3%	0.9%	0%	(-)	23%	38%		39%		19%	37%		33%	
Wei	2,197	3%	0.0%	n.a.		22%	25%	(-)	38%		20%	19%	(-)	33%	
Unw	2,075	9%	1.7%	0%	(-)	22%	53%		47%		19%	51%		41%	
Sep	2,075	9%	0.0%	n.a.		17%	33%		45%		15%	27%		40%	
Poo	2,197	3%	1.7%	0%	(-)	27%	42%		40%		24%	38%		34%	
One	1,505	0%	0.0%	n.a.		17%	50%		49%		14%	49%		41%	
Ad1	1,385	8%	1.5%	0%	(-)	28%	35%		40%		26%	29%		36%	
Ad2	1,382	9%	1.1%	0%	(-)	22%	33%		38%		20%	29%		35%	
L1	1,512	0%	2.5%	0%	(-)	24%	0%	(-)	18%	(-)	23%	0%	(-)	13%	(-)
L2	1,394	8%	0.0%	n.a.		18%	52%		51%		15%	56%		42%	
Ma	1,366	10%	0.1%	0%	(-)	25%	69%		61%		20%	60%		59%	
2Kat3	1,488	2%	0.2%	0%	(-)	22%	26%		34%		19%	16%	(-)	26%	
2Kat1	1,418	6%	2.4%	0%	(-)	22%	38%		45%		20%	34%		40%	
4Kat2	1,366	10%	0.0%	n.a.		23%	52%		49%		20%	52%		47%	

Table 56: Results of the extreme bounds analysis for the independent variable SingleAge aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a single-mover and age (as captured by the interaction effect SingleAge) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,272 dependent variables reveals a vague tendency for a negative sign of SingleAge which can be interpreted in two alternative ways: first, as a vague tendency for an additional negative single-mover effect when migrants turn one year older compared to non-single-movers; second, as a vague tendency for an additional negative age-effect for single-movers compared to non-single-movers.

4.3.11.1.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of SingleAge in explaining the probability of being risk-averse over the 4,272 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three findings are noteworthy: First, comparing results of subgroups referring to Head's individual versus family income, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to family income (Fam) is about 1.4 times (1.6 times) the percentage of robust relations in the subgroup referring to Head's individual income (Ind). Second, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to variance (Var) is about 1.4 times (1.9 times) the percentage of robust relations in the subgroup referring to semi-variance (LP2). Third, comparing the results of the subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.6 times (1.6 times) the percentage of robust relations in the subgroup referring to separate clustering in each year (Sep).

Tendency for a sign

The finding of a vague tendency for a negative sign of SingleAge for the 4,272 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different types of clustering (denoted by Sep and Poo), and not sensitive to (iv) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2).

However, the finding of a vague tendency for a negative sign of SingleAge is sensitive to (i) different risk-measures (denoted by Var and LP2), (ii) different education definitions (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iv) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (v) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.2 Interaction of single-moves and Edu2

4.3.11.2.1 Results irrespective of the way risk-attitudes are estimated

<u>Significance, robustness, and tendency for a sign</u>

The dummy variable SingleEdu2 captures the interaction effect of Head being a single-mover and having a high school diploma or an associate's degree as highest education level. The influence of SingleEdu2 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 3,915 dependent variables since quasi-complete separation occurs only for about 14% of the 4,536 dependent variables.[307] Consequently, I focus on the statistical analysis of SingleEdu2.

All statistical results on the influence of SingleEdu2 on the probability of being risk-averse are summarized in Table 57, p. 303. Based on the sample of 3,915 dependent variables, SingleEdu2 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, SingleEdu2 is also not significantly related to the probability of being risk-averse. Yet, I observe a vague tendency for a negative sign of SingleEdu2.

[307] Concerning the 621 dependent variables (about 14% of 4,536) where quasi-complete separation occurs, it is noteworthy that all single-mover with education level two (SingleEdu2=1) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	1,306	14%	1.5%	9%	(-)	40%	34%		51%		39%	33%		50%	
Ind	2,268	0%	1.8%	33%		36%	43%		54%		32%	49%		56%	
Fam	1,647	27%	1.4%	0%	(-)	46%	24%	(-)	45%		51%	20%	(-)	41%	
Ann	1,109	14%	2.6%	9%	(-)	40%	33%		49%		42%	31%		50%	
Wor	552	15%	0.3%	100%	(+)	40%	41%		56%		38%	39%		56%	
Lif	575	11%	0.5%	0%	(-)	47%	14%	(-)	36%		45%	16%	(-)	35%	
Var	575	11%	0.3%	0%	(-)	45%	22%	(-)	40%		44%	25%	(-)	37%	
LP2	552	15%	1.5%	20%	(-)	34%	47%		57%		31%	45%		59%	
Ed1	552	15%	3.1%	30%		36%	57%		64%		35%	55%		63%	
Ed2	2,089	8%	0.3%	100%	(+)	44%	43%		57%		39%	45%		59%	
Ed3	1,826	19%	2.9%	11%	(-)	36%	22%	(-)	42%		41%	20%	(-)	40%	
Ed4	1,923	15%	0.8%	68%		38%	40%		53%		40%	37%		53%	
Ed5	1,992	12%	2.3%	0%	(-)	43%	28%		48%		40%	29%		46%	
Ed6	1,344	11%	0.9%	100%	(+)	48%	47%		63%		36%	63%		64%	
Wei	1,285	15%	1.6%	0%	(-)	33%	25%	(-)	46%		38%	19%	(-)	44%	
Unw	1,286	15%	2.2%	0%	(-)	40%	24%	(-)	41%		45%	20%	(-)	40%	
Sep	1,512	0%	1.6%	25%	(-)	41%	45%		60%		43%	42%		60%	
Poo	1,299	14%	1.6%	0%	(-)	49%	27%		45%		46%	26%		43%	
One	1,104	27%	1.5%	30%		31%	27%		43%		28%	29%		44%	
Ad1	1,419	6%	0.6%	33%		33%	42%		55%		32%	42%		53%	
Ad2	1,312	13%	1.7%	12%	(-)	37%	35%		54%		35%	33%		54%	
L1	1,184	22%	2.4%	19%	(-)	53%	27%		40%		54%	26%		41%	
L2	1,306	14%	1.5%	9%	(-)	40%	34%		51%		39%	33%		50%	
Ma	2,268	0%	1.8%	33%		36%	43%		54%		32%	49%		56%	
2Kat3	1,647	27%	1.4%	0%	(-)	46%	24%	(-)	45%		51%	20%	(-)	41%	
2Kat1	1,109	14%	2.6%	9%	(-)	40%	33%		49%		42%	31%		50%	
4Kat2	552	15%	0.3%	100%	(+)	40%	41%		56%		38%	39%		56%	

Table 57: Results of the extreme bounds analysis for the independent variable SingleEdu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a single-mover and having a high school diploma or an associate's degree as highest education level (as captured by the interaction effect SingleEdu2) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 3,915 dependent variables reveals a vague tendency for a negative sign of SingleEdu2 which can be interpreted in two alternative ways: first, as a vague tendency for an additional negative single-mover effect for single-movers having a high school diploma or an associate's degree (Edu2=1) compared to non-single-movers and single-movers having other education levels (i.e., single-movers having less than a high school diploma (Edu1=1) and single-movers having a bachelor's degree or higher (Edu3=1)); second, as a vague tendency for an additional negative education-effect of having a high school diploma or an associate's degree (Edu2=1) for single-movers compared to non-single-movers and single-movers having other education levels (i.e., single-movers having less than a high school diploma (Edu1=1) and single-movers having a bachelor's degree or higher (Edu3=1)). Note that the vague tendency for negative sign of SingleEdu2 is in line with the remaining 621 dependent variables where quasi- complete separation occurs in the way that all single-movers with education level two are risk-averse.

4.3.11.2.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of SingleEdu2 in explaining the probability of being risk-averse over the 3,915 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 1.3 times (1.6 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-aversion are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a vague tendency for a negative sign of SingleEdu2 for the 3,915 dependent variables statistically investigated is sensitive to all competing solutions. In detail, the vague tendency for a negative sign of SingleEdu2 is sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.3 Interaction of single-moves and Edu3

4.3.11.3.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable SingleEdu3 captures the interaction effect of Head being a single-mover and having a bachelor's degree or higher as highest education level. The influence of SingleEdu3 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 3,758 dependent variables since quasi-complete separation occurs for about 17% of the 4,536 dependent variables.[308] Consequently, I focus on the statistical analysis of SingleEdu3.

All statistical results on the influence of SingleEdu3 on the probability of being risk-averse are summarized in Table 58, p. 306. Based on the sample of 3,758 dependent variables, SingleEdu3 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, SingleEdu3 is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a negative sign of SingleEdu3.

[308] Concerning the 778 dependent variables (about 17% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 750 out of 778 all single-movers with education level three (SingleEdu3=1) are risk-averse and (ii) for 28 out of 778 not all but 98% of the single-movers with education level three (SingleEdu3=1) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	3,758	17%	5%	0%	(-)	40%	10%	(-)	22%	(-)	37%	5%	(-)	21%	(-)
Ind	3,325	14%	5%	0%	(-)	41%	11%	(-)	22%	(-)	38%	6%	(-)	20%	(-)
Fam	433	33%	3%	0%	(-)	28%	7%	(-)	23%	(-)	29%	0%	(-)	24%	(-)
Ann	1,252	17%	5%	0%	(-)	40%	10%	(-)	21%	(-)	38%	5%	(-)	20%	(-)
Wor	1,254	17%	4%	0%	(-)	39%	11%	(-)	23%	(-)	37%	5%	(-)	22%	(-)
Lif	1,252	17%	5%	0%	(-)	40%	11%	(-)	21%	(-)	37%	5%	(-)	21%	(-)
Var	2,215	2%	5%	0%	(-)	38%	7%	(-)	22%	(-)	35%	4%	(-)	22%	(-)
LP2	1,543	32%	5%	0%	(-)	42%	15%	(-)	22%	(-)	40%	7%	(-)	19%	(-)
Ed1	967	25%	4%	0%	(-)	35%	10%	(-)	22%	(-)	33%	3%	(-)	21%	(-)
Ed2	531	18%	5%	0%	(-)	39%	12%	(-)	21%	(-)	38%	7%	(-)	18%	(-)
Ed3	527	19%	4%	0%	(-)	39%	8%	(-)	23%	(-)	35%	5%	(-)	21%	(-)
Ed4	527	19%	4%	0%	(-)	35%	10%	(-)	24%	(-)	33%	5%	(-)	25%	(-)
Ed5	603	7%	6%	0%	(-)	47%	11%	(-)	19%	(-)	43%	4%	(-)	20%	(-)
Ed6	603	7%	6%	0%	(-)	45%	11%	(-)	22%	(-)	44%	7%	(-)	19%	(-)
Wei	1,918	15%	6%	0%	(-)	39%	9%	(-)	16%	(-)	37%	6%	(-)	18%	(-)
Unw	1,840	19%	4%	0%	(-)	40%	12%	(-)	27%		38%	4%	(-)	23%	(-)
Sep	1,869	18%	3%	0%	(-)	32%	19%	(-)	29%		27%	11%	(-)	27%	
Poo	1,889	17%	7%	0%	(-)	47%	5%	(-)	15%	(-)	47%	2%	(-)	15%	(-)
One	1,470	3%	9%	0%	(-)	59%	6%	(-)	11%	(-)	54%	0%	(-)	10%	(-)
Ad1	1,154	24%	4%	0%	(-)	25%	14%	(-)	27%		25%	13%	(-)	26%	
Ad2	1,134	25%	2%	0%	(-)	30%	18%	(-)	31%		28%	11%	(-)	30%	
L1	1,310	13%	4%	0%	(-)	34%	9%	(-)	17%	(-)	30%	2%	(-)	19%	(-)
L2	1,156	24%	2%	0%	(-)	40%	4%	(-)	15%	(-)	38%	0%	(-)	14%	(-)
Ma	1,292	15%	9%	0%	(-)	45%	17%	(-)	33%		44%	11%	(-)	28%	
2Kat3	1,405	7%	2%	0%	(-)	38%	4%	(-)	16%	(-)	36%	0%	(-)	15%	(-)
2Kat1	1,276	16%	3%	0%	(-)	40%	5%	(-)	18%	(-)	37%	0%	(-)	13%	(-)
4Kat2	1,077	29%	10%	0%	(-)	41%	25%	(-)	33%		38%	17%	(-)	38%	

Table 58: Results of the extreme bounds analysis for the independent variable SingleEdu3 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a single-mover and having a bachelor's degree or higher as highest education level (as captured by the interaction effect SingleEdu3) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 3,758 dependent variables reveals a strong tendency for a negative sign of SingleEdu3 which can be interpreted in two alternative ways: first, as a strong tendency for an additional negative single-mover effect for single-movers having a bachelor's degree or higher (Edu3=1) compared to non-single-movers and single-movers with other education levels (i.e., single-movers having less than a high school diploma (Edu1=1) and single-movers having a high school diploma or an associate's degree (Edu2=1)); second, as a strong tendency for an additional negative education-effect of having a bachelor's degree or higher (Edu3=1) for single-movers compared to non-single-movers and single-movers with other education levels (i.e., single-movers having less than a high school diploma (Edu1=1) and single-movers having a high school diploma or an associate's degree (Edu2=1)). Note that the strong tendency for negative sign of SingleEdu3 is in line with 96.4% of the remaining 778 dependent variables where quasi- complete separation occurs in the way that all single-movers with education level three are risk-averse.

4.3.11.3.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of SingleEdu3 in explaining the probability of being risk-averse over the 3,758 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing results of subgroups referring to Head's individual versus family income, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to Head's individual income (Ind) is about 1.5 times (1.3 times) the percentage of robust relations in the subgroup referring to family income (Fam). Second, comparing the results of the subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.5 times (1.7 times) the percentage of robust relations in the subgroup referring to separate clustering (Sep).

Tendency for a sign

The finding of a strong tendency for a negative sign of SingleEdu3 for the 3,758 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, SingleEdu3 is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.4 Interaction of pair-moves and Age

4.3.11.4.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable PairAge captures the interaction effect of Head being a pair-mover and Head's age. The influence of PairAge on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,387 dependent variables since quasi-complete separation occurs only for about 3% of the 4,536 dependent variables.[309] Consequently, I focus on the statistical analysis of PairAge.

All statistical results on the influence of PairAge on the probability of being risk-averse are summarized in Table 59, p. 309. Based on the sample of 4,387 dependent variables, PairAge does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, PairAge is also not significantly related to the probability of being risk-averse. I do not observe a tendency for a sign of PairAge.

[309] Concerning the 149 dependent variables (about 3% of 4,536) where quasi-complete separation occurs, it is noteworthy that all pair-movers are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,387	3%	0.4%	100%	(+)	8%	31%		34%		5%	73%		40%	
Ind	3,742	4%	0.5%	100%	(+)	8%	28%		31%		5%	77%	(+)	38%	
Fam	645	0%	0.0%	n.a.		8%	43%		48%		8%	59%		56%	
Ann	1,465	3%	0.4%	100%	(+)	8%	34%		35%		6%	76%	(+)	41%	
Wor	1,461	3%	0.4%	100%	(+)	7%	30%		33%		5%	71%		40%	
Lif	1,461	3%	0.4%	100%	(+)	8%	28%		34%		5%	71%		41%	
Var	2,233	2%	0.1%	100%	(+)	7%	51%		49%		6%	100%	(+)	54%	
LP2	2,154	5%	0.7%	100%	(+)	9%	15%	(-)	18%	(-)	4%	32%		26%	
Ed1	1,270	2%	0.5%	100%	(+)	8%	40%		41%		7%	66%		50%	
Ed2	625	4%	0.5%	100%	(+)	7%	25%	(-)	35%		6%	57%		42%	
Ed3	621	4%	0.5%	100%	(+)	9%	24%	(-)	24%	(-)	4%	87%	(+)	30%	
Ed4	621	4%	0.5%	100%	(+)	10%	30%		24%	(-)	4%	89%	(+)	31%	
Ed5	625	4%	0.5%	100%	(+)	8%	29%		36%		5%	81%	(+)	40%	
Ed6	625	4%	0.0%	n.a.		7%	24%	(-)	36%		4%	79%	(+)	40%	
Wei	2,154	5%	0.7%	100%	(+)	7%	40%		37%		4%	76%	(+)	44%	
Unw	2,233	2%	0.1%	100%	(+)	9%	24%	(-)	31%		7%	72%		37%	
Sep	2,220	2%	0.0%	n.a.		6%	21%	(-)	37%		2%	67%		43%	
Poo	2,167	4%	0.8%	100%	(+)	10%	36%		31%		9%	74%		37%	
One	1,512	0%	1.0%	100%	(+)	7%	30%		28%		2%	97%	(+)	33%	
Ad1	1,441	5%	0.2%	100%	(+)	10%	35%		38%		8%	72%		44%	
Ad2	1,434	5%	0.0%	n.a.		8%	25%	(-)	36%		6%	66%		43%	
L1	1,512	0%	0.0%	n.a.		0%	n.a.	n.a.	48%		2%	100%	(+)	51%	
L2	1,512	0%	1.2%	100%	(+)	11%	18%	(-)	19%	(-)	4%	45%		28%	
Ma	1,363	10%	0.0%	n.a.		13%	43%		34%		10%	81%	(+)	42%	
2Kat3	1,512	0%	0.0%	n.a.		3%	0%	(-)	28%		0%	n.a.	n.a.	34%	
2Kat1	1,470	3%	0.0%	n.a.		5%	24%	(-)	33%		5%	92%	(+)	37%	
4Kat2	1,405	7%	1.2%	100%	(+)	17%	38%		41%		11%	63%		51%	

Table 59: Results of the extreme bounds analysis for the independent variable PairAge aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a pair-mover and age (as captured by the interaction effect PairAge) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,387 dependent variables reveals a no tendency for a sign of PairAge which can be interpreted in two alternative ways: first, as no additional pair-mover effect when pair-movers turn one year older compared to non-pair-movers; second, as no additional age-effect for pair-movers compared to non-pair-movers.

4.3.11.4.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of PairAge in explaining the probability of being risk-averse over the 4,387 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, two findings are noteworthy: First, comparing results of subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.7 times (4.5 times) the percentage of robust relations in the subgroup referring to separate clustering in each year (Sep). Second, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-aversion are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of no tendency for a sign of PairAge for the 4,387 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different education definitions (denoted by Ed1 to Ed6), (iv) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (v) different types of clustering (denoted by Sep and Poo), and not sensitive to (vi) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2).

However, the finding of no tendency for a sign of PairAge is sensitive to (i) different risk-measures (denoted by Var and LP2), (ii) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (iii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.5 Interaction of pair-moves and Edu2

4.3.11.5.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable PairEdu2 captures the interaction effect of Head being a pair-mover and having a high school diploma or an associate's degree as highest education level. The influence of PairEdu2 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,357 dependent variables since quasi-complete separation occurs only for about 4% of the 4,536 dependent variables.[310] Consequently, I focus on the statistical analysis of PairEdu2.

All statistical results on the influence of PairEdu2 on the probability of being risk-averse are summarized in Table 60, p. 312. Based on the sample of 4,357 dependent variables, PairEdu2 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, PairEdu2 is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a negative sign of PairEdu2.

[310] Concerning the 179 dependent variables (about 4% of 4,536) where quasi-complete separation occurs, it is noteworthy that all pair-movers are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	**number of dependent variables**	**missing values**	**Leamer**			**Sala-I-Martin weighted**					**Sala-I-Martin unweighted**				
			significant	**positive**		**CDF(0) ≥95%**	**positive**		**positive**		**CDF(0) ≥95%**	**positive**		**positive**	
	n	**% of 4,536**	**% of n**	**% of significant**		**% of n**	**% of robust**		**% of n**		**% of n**	**% of robust**		**% of n**	
Total	4,357	4%	19%	0%	(-)	55%	13%	(-)	29%		61%	6%	(-)	20%	(-)
Ind	3,712	5%	20%	0%	(-)	57%	13%	(-)	29%		63%	6%	(-)	19%	(-)
Fam	645	0%	10%	0%	(-)	44%	13%	(-)	29%		52%	4%	(-)	25%	(-)
Ann	1,455	4%	19%	0%	(-)	55%	13%	(-)	29%		62%	6%	(-)	20%	(-)
Wor	1,451	4%	18%	0%	(-)	55%	13%	(-)	29%		62%	6%	(-)	19%	(-)
Lif	1,451	4%	19%	0%	(-)	55%	13%	(-)	29%		61%	6%	(-)	20%	(-)
Var	2,203	3%	22%	0%	(-)	47%	17%	(-)	40%		52%	5%	(-)	24%	(-)
LP2	2,154	5%	16%	0%	(-)	63%	10%	(-)	17%	(-)	71%	6%	(-)	15%	(-)
Ed1	1,264	2%	14%	0%	(-)	52%	13%	(-)	30%		58%	5%	(-)	21%	(-)
Ed2	619	4%	17%	0%	(-)	57%	13%	(-)	29%		62%	5%	(-)	18%	(-)
Ed3	618	5%	24%	0%	(-)	53%	11%	(-)	29%		61%	5%	(-)	22%	(-)
Ed4	618	5%	21%	0%	(-)	57%	16%	(-)	33%		62%	5%	(-)	23%	(-)
Ed5	619	4%	19%	0%	(-)	59%	14%	(-)	26%		66%	8%	(-)	16%	(-)
Ed6	619	4%	19%	0%	(-)	56%	12%	(-)	26%		63%	6%	(-)	16%	(-)
Wei	2,154	5%	17%	0%	(-)	66%	13%	(-)	18%	(-)	74%	3%	(-)	12%	(-)
Unw	2,203	3%	20%	0%	(-)	44%	14%	(-)	40%		49%	9%	(-)	27%	
Sep	2,190	3%	13%	0%	(-)	46%	13%	(-)	30%		57%	7%	(-)	22%	(-)
Poo	2,167	4%	24%	0%	(-)	64%	14%	(-)	28%		66%	4%	(-)	17%	(-)
One	1,512	0%	9%	0%	(-)	60%	23%	(-)	39%		54%	7%	(-)	21%	(-)
Ad1	1,429	5%	25%	0%	(-)	52%	7%	(-)	24%	(-)	66%	5%	(-)	18%	(-)
Ad2	1,416	6%	22%	0%	(-)	53%	8%	(-)	23%	(-)	64%	5%	(-)	19%	(-)
L1	1,512	0%	19%	0%	(-)	54%	8%	(-)	30%		60%	4%	(-)	21%	(-)
L2	1,512	0%	10%	0%	(-)	56%	13%	(-)	29%		59%	5%	(-)	22%	(-)
Ma	1,333	12%	27%	0%	(-)	56%	19%	(-)	28%		66%	8%	(-)	15%	(-)
2Kat3	1,512	0%	12%	0%	(-)	29%	14%	(-)	30%		49%	4%	(-)	18%	(-)
2Kat1	1,470	3%	23%	0%	(-)	61%	15%	(-)	31%		62%	7%	(-)	19%	(-)
4Kat2	1,375	9%	20%	0%	(-)	77%	12%	(-)	26%		74%	5%	(-)	22%	(-)

Table 60: Results of the extreme bounds analysis for the independent variable PairEdu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a pair-movers and having a high school diploma or an associate's degree as highest education level (as captured by the interaction effect PairEdu2) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,357 dependent variables reveals a strong tendency for a negative sign of PairEdu2 which can be interpreted in two alternative ways: first, as a strong tendency for an additional negative pair-mover effect for pair-movers having a high school diploma or an associate's degree (Edu2=1) compared to non-pair-movers and pair-movers having other education levels (i.e., pair-movers having less than a high school diploma (Edu1=1) and pair-movers having a bachelor's degree or higher (Edu3=1)); second, as a strong tendency for an additional negative education-effect of having a high school diploma or an associate's degree (Edu2=1) for pair-movers compared to non-pair-movers and pair-movers having other education levels (i.e., pair-movers having less than a high school diploma (Edu1=1) and pair-movers having a bachelor's degree or higher (Edu3=1)).

4.3.11.5.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of PairEdu2 in explaining the probability of being risk-averse over the 4,357 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, four finding are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 1.3 times (1.4 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to weighted samples (Wei) is about 1.5 times (1.5 times) the percentage of robust relations in the subgroup referring to unweighted samples (Unw). Third, comparing results of subgroups referring to different types of clustering, I find the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.4 times (1.2 times) the percentage of robust relations in the subgroup referring to separate clustering in each year (Sep). Fourth, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-aversion are deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by

Sala-I-Martin's weighted and unweighted criteria increase visibly. That is, in the subgroup that only includes migrants with the most pronounced degree of risk-attitude (4Kat2) Sala-I-Martin's weighted and unweighted criteria almost meet the 75%-criterion of robustness with 77% and 74% of the relations being robust.

Tendency for a sign

The finding of a strong tendency for a negative sign of PairEdu2 for the 4,357 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, PairEdu2 is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.6 Interaction of pair-moves and Edu3

4.3.11.6.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable PairEdu3 captures the interaction effect of Head being a pair-mover and having a bachelor's degree or higher as highest education level. The influence of PairEdu3 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 1,701 dependent variables although quasi-complete separation occurs for notable 63% of the 4,536 dependent variables.[311]

[311] Concerning the 2,835 dependent variables (about 63% of 4,536) where quasi-complete separation occurs, it is noteworthy that all pair-movers with education level three (PairEdu3=1) are risk-averse although there are enough pair-movers with this education level in the sample, i.e., 31 to 32 of the 321 migrants in my sample are pair-movers with education level three.

All statistical results on the influence of PairEdu3 on the probability of being risk-averse are summarized in Table 61, p. 316. Based on the sample of 1,701 dependent variables, PairEdu3 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, PairEdu3 is also not significantly related to the probability of being risk-averse. I also do not observe a tendency for a sign of PairEdu3.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	1,701	63%	0.8%	0%	(-)	12%	56%		57%		24%	10%	(-)	44%	
Ind	1,326	66%	0.6%	0%	(-)	13%	63%		58%		23%	11%	(-)	48%	
Fam	375	42%	1.7%	0%	(-)	10%	26%		52%		26%	6%	(-)	31%	
Ann	561	63%	0.7%	0%	(-)	12%	57%		56%		24%	10%	(-)	45%	
Wor	574	62%	0.8%	0%	(-)	12%	55%		57%		24%	10%	(-)	43%	
Lif	566	63%	0.9%	0%	(-)	12%	57%		57%		25%	9%	(-)	44%	
Var	806	64%	0.4%	0%	(-)	9%	43%		49%		33%	7%	(-)	36%	
LP2	895	61%	1.1%	0%	(-)	16%	63%		64%		16%	15%	(-)	51%	
Ed1	595	54%	1.1%	0%	(-)	11%	44%		53%		25%	6%	(-)	36%	
Ed2	220	66%	1.4%	0%	(-)	14%	60%		54%		30%	14%	(-)	46%	
Ed3	223	66%	0.0%	n.a.		13%	60%		49%		17%	8%	(-)	40%	
Ed4	223	66%	0.0%	n.a.		13%	60%		49%		21%	6%	(-)	42%	
Ed5	220	66%	0.5%	0%	(-)	13%	68%		71%		21%	13%	(-)	55%	
Ed6	220	66%	1.5%	0%	(-)	11%	63%		69%		27%	15%	(-)	57%	
Wei	791	65%	0.3%	0%	(-)	7%	63%		53%		22%	14%	(-)	29%	
Unw	910	60%	1.3%	0%	(-)	17%	54%		60%		25%	6%	(-)	57%	
Sep	1,260	44%	1.3%	0%	(-)	13%	60%		63%		22%	14%	(-)	51%	
Poo	441	81%	0.3%	0%	(-)	10%	41%		39%		29%	0%	(-)	24%	(-)
One	284	81%	0.0%	n.a.		13%	92%	(+)	92%	(+)	8%	82%	(+)	78%	(+)
Ad1	700	54%	0.8%	0%	(-)	10%	60%		48%		28%	5%	(-)	36%	
Ad2	717	53%	1.6%	0%	(-)	14%	41%		51%		26%	6%	(-)	38%	
L1	669	56%	0.1%	0%	(-)	7%	72%		63%		19%	7%	(-)	44%	
L2	527	65%	0.0%	n.a.		10%	92%	(+)	65%		6%	36%		51%	
Ma	505	67%	2.2%	0%	(-)	21%	31%		40%		50%	7%	(-)	37%	
2Kat3	1,019	33%	0.4%	0%	(-)	7%	47%		60%		22%	8%	(-)	43%	
2Kat1	440	71%	0.0%	n.a.		14%	90%	(+)	67%		18%	12%	(-)	59%	
4Kat2	242	84%	2.0%	0%	(-)	32%	38%		25%	(-)	43%	12%	(-)	22%	(-)

Table 61: Results of the extreme bounds analysis for the independent variable PairEdu3 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

<u>Comparison to the literature and interpretation</u>

Since the interaction of being a pair-mover and having a bachelor's degree or higher as highest education level (as captured by the interaction effect PairEdu3) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 1,701 dependent variables reveals no tendency for a sign of PairEdu3 which can be interpreted in two alternative ways: first, as no additional pair-mover effect for pair-movers having a bachelor's degree or higher (Edu3=1) compared to non-pair-movers and pair-movers having other education levels (i.e., pair-movers having less than a high school diploma (Edu1=1) and pair-movers having a high school diploma or an associate's degree (Edu2=1)); second, as no additional education-effect of having a bachelor's degree or higher (Edu3=1) for pair-movers compared to non-pair-movers and pair-movers having other education levels (i.e., pair-movers having less than a high school diploma (Edu1=1) and pair-movers having a high school diploma or an associate's degree (Edu2=1)).

Nevertheless, statistical statements of non-robustness and no tendency for sign of PairEdu3 must be interpreted with caution for two reasons: First, these statements are not in line with the majority of dependent variables (2,835 dependent variables or about 63% of 4,536, respectively) where quasi-complete separation occurs in the way that all pair-movers with education level three (i.e., having a bachelor's degree or higher, are risk-averse) are risk-averse. Second, for the dependent variables where quasi-complete separation occurs there are not only a few but considerably 31 to 32 out of 321 migrants in the sample that are pair-movers and have education level three (PairEdu3=1).

4.3.11.6.2 Sensitivity for subgroups of dependent variables

<u>Robustness</u>

The fragility and non-significance of PairEdu3 in explaining the probability of being risk-averse over the 1,701 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, one finding is noteworthy: Comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to unweighted samples (Unw) is about 2.4 times (1.4 times) the percentage of robust relations in the subgroup referring to weighted samples (Wei).

Tendency for a sign

The finding of no tendency for a sign of PairEdu3 for the 1,701 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different education definitions (denoted by Ed1 to Ed6), (iv) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), and not sensitive to (v) different types of clustering (denoted by Sep and Poo).

However, the finding of no tendency for a sign of PairEdu3 is sensitive to (i) different risk-measures (denoted by Var and LP2), (ii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (iii) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (iv) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.7 Interaction of family-moves and Age

4.3.11.7.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable FamilyAge captures the interaction effect of Head being a family-mover and Head's age. The influence of FamilyAge on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,526 dependent variables since quasi-complete separation occurs only for about 0.2% of the 4,536 dependent variables.[312] Consequently, I focus on the statistical analysis of FamilyAge.

All statistical results on the influence of FamilyAge on the probability of being risk-averse are summarized in Table 62, p. 319. Based on the sample of 4,526 dependent variables, FamilyAge does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, FamilyAge is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a positive sign of FamilyAge.

[312] Concerning the 10 dependent variables (about 0.2% of 4,536) where quasi-complete separation occurs, it is noteworthy that all family-movers are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,526	0.2%	0.8%	47%		23%	84%	(+)	80%	(+)	23%	79%	(+)	79%	(+)
Ind	3,878	0.3%	1.0%	47%		22%	85%	(+)	81%	(+)	22%	82%	(+)	79%	(+)
Fam	648	0.0%	0.0%	n.a.		28%	77%	(+)	76%	(+)	29%	70%		74%	
Ann	1,506	0.4%	0.9%	46%		22%	83%	(+)	79%	(+)	22%	80%	(+)	78%	(+)
Wor	1,510	0.1%	0.9%	46%		23%	83%	(+)	80%	(+)	23%	79%	(+)	79%	(+)
Lif	1,510	0.1%	0.8%	50%		23%	84%	(+)	80%	(+)	23%	79%	(+)	79%	(+)
Var	2,268	0.0%	1.4%	56%		29%	92%	(+)	82%	(+)	30%	90%	(+)	81%	(+)
LP2	2,258	0.4%	0.3%	0%	(-)	17%	69%		78%	(+)	16%	60%		76%	(+)
Ed1	1,293	0.2%	0.4%	60%		26%	81%	(+)	78%	(+)	27%	76%	(+)	77%	(+)
Ed2	645	0.5%	0.5%	0%	(-)	22%	86%	(+)	82%	(+)	22%	81%	(+)	81%	(+)
Ed3	647	0.2%	1.9%	75%	(+)	20%	93%	(+)	78%	(+)	19%	88%	(+)	77%	(+)
Ed4	647	0.2%	0.9%	0%	(-)	15%	81%	(+)	78%	(+)	15%	75%	(+)	76%	(+)
Ed5	647	0.2%	1.4%	67%		27%	83%	(+)	83%	(+)	26%	81%	(+)	81%	(+)
Ed6	647	0.2%	0.5%	0%	(-)	26%	82%	(+)	82%	(+)	25%	79%	(+)	81%	(+)
Wei	2,258	0.4%	1.1%	75%	(+)	28%	89%	(+)	80%	(+)	26%	89%	(+)	82%	(+)
Unw	2,268	0.0%	0.6%	0%	(-)	18%	76%	(+)	80%	(+)	19%	66%		75%	(+)
Sep	2,268	0.0%	0.7%	100%	(+)	20%	86%	(+)	79%	(+)	19%	80%	(+)	77%	(+)
Poo	2,258	0.4%	1.0%	13%	(-)	26%	82%	(+)	81%	(+)	26%	79%	(+)	81%	(+)
One	1,512	0.0%	0.9%	0%	(-)	20%	74%		71%		20%	61%		72%	
Ad1	1,508	0.3%	1.0%	60%		26%	87%	(+)	83%	(+)	25%	88%	(+)	82%	(+)
Ad2	1,506	0.4%	0.6%	100%	(+)	23%	87%	(+)	86%	(+)	23%	86%	(+)	82%	(+)
L1	1,512	0.0%	0.0%	n.a.		34%	99%	(+)	95%	(+)	34%	99%	(+)	96%	(+)
L2	1,512	0.0%	1.2%	100%	(+)	16%	90%	(+)	89%	(+)	14%	89%	(+)	87%	(+)
Ma	1,502	0.7%	1.3%	0%	(-)	19%	50%		55%		20%	40%		53%	
2Kat3	1,512	0.0%	1.2%	100%	(+)	31%	96%	(+)	87%	(+)	31%	96%	(+)	87%	(+)
2Kat1	1,502	0.7%	0.9%	0%	(-)	22%	86%	(+)	82%	(+)	24%	81%	(+)	83%	(+)
4Kat2	1,512	0.0%	0.4%	0%	(-)	16%	55%		71%		13%	39%		66%	

Table 62: Results of the extreme bounds analysis for the independent variable FamilyAge aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of age and being a family-mover (as captured by the interaction effect FamilyAge) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,526 dependent variables reveals a strong tendency for a positive sign of FamilyAge which can be interpreted in two alternative ways: first, as a strong tendency for an additional positive family-mover effect when family-movers turn one year older compared to non-famiy-movers; second, as a strong tendency for an additional positive age-effect for family-movers compared to non-family-movers.

4.3.11.7.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of FamilyAge in explaining the probability of being risk-averse over the 4,526 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three findings are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to variance (Var) is about 1.7 times (1.9 times) the percentage of robust relations in the subgroup referring to semi-variance (LP2). Second, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to weighted samples (Wei) is about 1.6 times (1.4 times) the percentage of robust relations in the subgroup referring to unweighted samples (Unw). Third, comparing results of subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.3 times (1.4 times) the percentage of robust relations in the subgroup referring to separate clustering in each year (Sep).

Tendency for a sign

The finding of a strong tendency for a positive sign of FamilyAge for the 4,526 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), and not sensitive (vi) different types of clustering (denoted by Sep and Poo).

However, the finding of a strong tendency for a positive sign of FamilyAge is sensitive to (i) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (ii) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (iii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.8 Interaction of family-moves and Edu2

4.3.11.8.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable FamilyEdu2 captures the interaction effect of Head being a family-mover and having a high school diploma or an associate's degree as highest education level. The influence of FamilyEdu2 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,513 dependent variables since quasi-complete separation occurs only for about 1% of the 4,536 dependent variables.[313] Consequently, I focus on the statistical analysis of FamilyEdu2.

All statistical results on the influence of FamilyEdu2 on the probability of being risk-averse are summarized in Table 63, p. 322. Based on the sample of 4,513 dependent variables, FamilyEdu2 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, FamilyEdu2 is also not significantly related to the probability of being risk-averse. Yet, I observe a vague tendency for a positive sign of FamilyEdu2.

[313] Concerning the 23 dependent variables (about 1% of 4,536) where quasi-complete separation occurs, it is noteworthy that the situation of quasi-complete separation most certainly is a result of an extremely small number of observations for risk-seeking migrants for these dependent variables, i.e., there are no more than 5 out of 321 migrants that are risk-seeking. This results in two different types of quasi-complete separation for the 23 dependent variables where quasi-complete separation occurs: (i) for 13 out of 23 all family-movers with education level two (FamilyEdu2=1) are risk-averse, (ii), for 4 out of 23 not all but 98% of the family-movers with education level two (FamilyEdu2=1) are risk-averse, and (iii) for 6 out of the 23 all non-family-movers (FamilyEdu2=0) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,513	1%	1.6%	52%		29%	66%		62%		32%	53%		57%	
Ind	3,875	0%	1.8%	54%		29%	68%		61%		32%	52%		56%	
Fam	638	2%	0.3%	0%	(-)	27%	58%		64%		29%	56%		60%	
Ann	1,501	1%	1.6%	54%		29%	67%		62%		31%	54%		57%	
Wor	1,506	0%	1.3%	60%		29%	65%		61%		32%	52%		56%	
Lif	1,506	0%	1.9%	45%		29%	66%		62%		32%	52%		56%	
Var	2,260	0%	1.5%	0%	(-)	19%	37%		45%		26%	23%	(-)	41%	
LP2	2,253	1%	1.7%	100%	(+)	39%	81%	(+)	78%	(+)	37%	74%		72%	
Ed1	1,283	1%	1.9%	63%		27%	60%		62%		32%	51%		59%	
Ed2	645	0%	2.0%	46%		31%	61%		56%		33%	49%		53%	
Ed3	647	0%	1.4%	67%		30%	82%	(+)	74%		30%	72%		67%	
Ed4	647	0%	1.4%	67%		32%	72%		67%		35%	58%		61%	
Ed5	647	0%	1.5%	30%		25%	68%		58%		30%	44%		51%	
Ed6	644	1%	1.2%	25%	(-)	30%	60%		52%		31%	45%		47%	
Wei	2,253	1%	1.3%	0%	(-)	34%	60%		62%		36%	48%		57%	
Unw	2,260	0%	1.9%	88%	(+)	23%	75%	(+)	61%		27%	59%		56%	
Sep	2,266	0%	0.2%	0%	(-)	23%	66%		61%		26%	46%		55%	
Poo	2,247	1%	3.0%	56%		35%	67%		63%		37%	58%		58%	
One	1,504	1%	2.3%	0%	(-)	30%	29%		39%		35%	16%	(-)	33%	
Ad1	1,508	0%	1.0%	100%	(+)	24%	89%	(+)	73%		27%	75%	(+)	69%	
Ad2	1,501	1%	1.5%	100%	(+)	33%	85%	(+)	73%		33%	74%		68%	
L1	1,512	0%	2.5%	63%		31%	63%		52%		37%	50%		49%	
L2	1,509	0%	2.0%	47%		30%	82%	(+)	69%		32%	70%		66%	
Ma	1,492	1%	0.3%	0%	(-)	26%	52%		63%		27%	36%		54%	
2Kat3	1,512	0%	0.0%	n.a.		20%	66%		55%		22%	49%		53%	
2Kat1	1,502	1%	0.4%	100%	(+)	23%	60%		61%		28%	45%		54%	
4Kat2	1,499	1%	4.4%	48%		44%	70%		68%		45%	59%		62%	

Table 63: Results of the extreme bounds analysis for the independent variable FamilyEdu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a family-mover and having a high school diploma or an associate's degree as highest education level (as captured by the interaction effect FamilyEdu2) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,513 dependent variables reveals a vague tendency for a positive sign of FamilyEdu2 which can be interpreted in two alternative ways: first, as a vague tendency for an additional positive family-mover effect for family-movers having a high school diploma or an associate's degree (Edu2=1) compared to non-family-movers and family-movers having other education levels (i.e., family-movers having less than a high school diploma (Edu1=1) and family-movers having a bachelor's degree or higher (Edu3=1)); second, as a vague tendency for an additional positive education-effect of having a high school diploma or an associate's degree (Edu2=1) for family-movers compared to non-family-movers and family-movers having other education levels (i.e., family-movers having less than a high school diploma (Edu1=1) and family-movers having a bachelor's degree or higher (Edu3=1)).

4.3.11.8.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of FamilyEdu2 in explaining the probability of being risk-averse over the 4,513 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, four findings are noteworthy: First, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 2.1 times (1.4 times) the percentage of robust relations in the subgroup referring to variance (Var). Second, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to weighted samples (Wei) is about 1.5 times (1.3 times) the percentage of robust relations in the subgroup referring to unweighted samples (Unw). Third, comparing the results of the subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.5 times (1.4 times) the percentage of robust relations in the subgroup referring to separate clustering in each year (Sep). Fourth, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-aversion are

deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a vague tendency for a positive sign of FamilyEdu2 for the 4,513 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam) and not sensitive to (ii) variations of the planning period (denoted by Ann, Wor, and Lif).

However, the finding of vague tendency for a positive sign of FamilyEdu2 is sensitive to (i) different risk-measures (denoted by Var and LP2), (ii) different education definitions (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iv) different types of clustering (denoted by Sep and Poo), (v) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (vi) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (vii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.11.9 Interaction of family-moves and Edu3

4.3.11.9.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable FamilyEdu3 captures the interaction effect of Head being a family-mover and having a bachelor's degree or higher as highest education level. The influence of FamilyEdu3 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,221 dependent variables since quasi-complete separation occurs only for about 7% of the 4,536 dependent variables.[314] Consequently, I focus on the statistical analysis of FamilyEdu3.

All statistical results on the influence of FamilyEdu3 on the probability of being risk-averse are summarized in Table 64, p. 325. Based on the sample of 4,221 dependent variables, FamilyEdu3 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, FamilyEdu3 is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a negative sign of FamilyEdu3.

[314] Concerning the 315 dependent variables (about 7% of 4,536) where quasi-complete separation occurs, it is noteworthy that all family-movers with education level three (FamilyEdu3=1) are risk-averse.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,221	7%	7%	0%	(-)	21%	24%	(-)	35%		24%	11%	(-)	35%	
Ind	3,633	7%	7%	0%	(-)	22%	23%	(-)	34%		25%	11%	(-)	34%	
Fam	588	9%	5%	0%	(-)	12%	24%	(-)	37%		19%	11%	(-)	40%	
Ann	1,402	7%	6%	0%	(-)	21%	23%	(-)	35%		23%	11%	(-)	35%	
Wor	1,416	6%	6%	0%	(-)	20%	24%	(-)	35%		24%	10%	(-)	34%	
Lif	1,403	7%	7%	0%	(-)	21%	24%	(-)	34%		24%	11%	(-)	34%	
Var	2,157	5%	2%	0%	(-)	13%	26%		42%		18%	14%	(-)	43%	
LP2	2,064	9%	11%	0%	(-)	30%	22%	(-)	26%		30%	9%	(-)	25%	(-)
Ed1	1,191	8%	7%	0%	(-)	18%	19%	(-)	34%		23%	11%	(-)	35%	
Ed2	604	7%	9%	0%	(-)	23%	17%	(-)	29%		26%	8%	(-)	31%	
Ed3	608	6%	9%	0%	(-)	24%	36%		33%		25%	13%	(-)	35%	
Ed4	611	6%	8%	0%	(-)	23%	27%		30%		27%	7%	(-)	28%	
Ed5	603	7%	3%	0%	(-)	19%	24%	(-)	41%		21%	17%	(-)	39%	
Ed6	604	7%	3%	0%	(-)	21%	19%	(-)	41%		21%	9%	(-)	38%	
Wei	2,037	10%	7%	0%	(-)	22%	20%	(-)	38%		24%	11%	(-)	35%	
Unw	2,184	4%	7%	0%	(-)	20%	27%		31%		24%	10%	(-)	34%	
Sep	2,161	5%	4%	0%	(-)	15%	17%	(-)	31%		21%	5%	(-)	31%	
Poo	2,060	9%	9%	0%	(-)	27%	27%		38%		27%	15%	(-)	38%	
One	1,381	9%	3%	0%	(-)	19%	25%	(-)	41%		21%	14%	(-)	36%	
Ad1	1,423	6%	8%	0%	(-)	19%	22%	(-)	32%		24%	11%	(-)	36%	
Ad2	1,417	6%	9%	0%	(-)	25%	24%	(-)	30%		26%	8%	(-)	32%	
L1	1,451	4%	8%	0%	(-)	20%	29%		39%		24%	2%	(-)	39%	
L2	1,475	2%	10%	0%	(-)	24%	10%	(-)	28%		26%	9%	(-)	28%	
Ma	1,295	14%	2%	0%	(-)	18%	37%		37%		20%	25%	(-)	37%	
2Kat3	1,470	3%	0%	n.a.		13%	64%		56%		9%	48%		53%	
2Kat1	1,462	3%	5%	0%	(-)	15%	21%	(-)	32%		16%	12%	(-)	36%	
4Kat2	1,289	15%	14%	0%	(-)	37%	8%	(-)	14%	(-)	49%	2%	(-)	12%	(-)

Table 64: Results of the extreme bounds analysis for the independent variable FamilyEdu3 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being a family-mover and having a bachelor's degree or higher as highest education level (as captured by the interaction effect FamilyEdu3) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,221 dependent variables reveals a strong tendency for a negative sign of FamilyEdu3 which can be interpreted in two alternative ways: first, as a strong tendency for an additional negative family-mover effect for family-movers having a bachelor's degree or higher (Edu3=1) compared to non-family-movers and family-movers having other education levels (i.e., family-movers having less than a high school diploma (Edu1=1) and family-movers having a high school diploma or an associate's degree (Edu2=1)); second, as a strong tendency for an additional negative education-effect of having a bachelor's degree or higher (Edu3=1) for family-movers compared to non-family-movers and family-movers having other education levels (i.e., family-movers having less than a high school diploma (Edu1=1) and family-movers having a high school diploma or an associate's degree (Edu2=1)).

4.3.11.9.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of FamilyEdu3 in explaining the probability of being risk-averse over the 4,221 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, four findings are noteworthy: First, comparing results of subgroups referring to Head's individual versus family income, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to Head's individual income (Ind) is about 1.8 times (1.3 times) the percentage of robust relations in the subgroup referring to family income (Fam). Second, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 2.3 times (1.7 times) the percentage of robust relations in the subgroup referring to variance (Var). Third, comparing the results of the subgroups referring to different types of clustering, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to pooled clustering (Poo) is about 1.8 times (1.3 times) the percentage of robust relations in the subgroup referring to separate clustering in each year (Sep). Fourth, comparing results of subgroups referring to different transformation rules, I find that when migrants with less pronounced degrees of risk-aversion are

deleted from the sample step by step (2Kat3, 2Kat1, 4Kat2), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a strong tendency for a negative sign of FamilyEdu3 for the 4,221 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), and not sensitive to (viii) different measurements of predictive errors (denoted by L1, L2, and Ma).

However, the finding of a strong tendency for a negative sign of FamilyEdu3 is sensitive to (i) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.12 Interactions with Divorce and their relation to risk-attitudes

In Part C, Section 4.3.8 I found that divorced Heads tend to possess a higher probability of being risk-averse than all other migrants i.e., Divorce exhibits a strong tendency for a positive sign. In this section I investigate whether these effects depend on other explanatory variables, namely (i) Head's age as captured by interaction effects with Age and (ii) Head's education level as captured by interaction effects with Edu2 and Edu3, respectively.

4.3.12.1 Interaction of Divorce and Age

4.3.12.1.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable DivorceAge captures the interaction effect of Head being divorced and Head's age. For about 64% of the 4,536 dependent variables in my study DivorceAge shows quasi-complete separation.[315] The fact that DivorceAge almost perfectly predicts risk-attitudes for about 64% of the

[315] Concerning the 2,891 dependent variables (about 64% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 2,888 out of 2,891 all divorced Heads are risk-averse and (ii) for 3 out of 2,891 only one divorced Head exists and this Head is risk-seeking.

dependent variables should not be overrated since there are no more than only 12 out of 321 migrants in my sample that are divorced. Therefore, I focus on the statistical analysis of DivorceAge based on a still sufficiently high sample of 1,645 dependent variables for which no quasi-complete separation occurs.

All statistical results on the influence of DivorceAge on the probability of being risk-averse are summarized in Table 65, p. 329. Based on 1,645 dependent variables, DivorceAge does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, DivorceAge is also not significantly related to the probability of being risk-averse. Yet, I observe a vague tendency for a positive sign of DivorceAge.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	1,645	64%	1.3%	15%	(-)	25%	52%		61%		24%	52%		61%	
Ind	1,512	61%	1.5%	15%	(-)	19%	67%		65%		18%	68%		65%	
Fam	133	79%	0.0%	n.a.		85%	13%	(-)	20%	(-)	89%	13%	(-)	20%	(-)
Ann	542	64%	1.6%	13%	(-)	24%	51%		61%		24%	51%		62%	
Wor	553	63%	0.8%	25%	(-)	24%	55%		61%		24%	55%		61%	
Lif	550	64%	1.6%	13%	(-)	25%	49%		61%		25%	49%		61%	
Var	1,048	54%	1.2%	11%	(-)	26%	52%		51%		26%	49%		52%	
LP2	597	74%	1.5%	18%	(-)	23%	51%		78%	(+)	20%	56%		78%	(+)
Ed1	386	70%	0.3%	0%	(-)	41%	33%		51%		40%	30%		51%	
Ed2	253	61%	1.1%	0%	(-)	17%	75%	(+)	66%		16%	80%	(+)	64%	
Ed3	249	62%	1.5%	30%		20%	65%		65%		19%	69%		67%	
Ed4	248	62%	2.0%	46%		21%	65%		65%		21%	65%		67%	
Ed5	256	60%	2.0%	0%	(-)	20%	57%		63%		19%	60%		63%	
Ed6	253	61%	2.0%	0%	(-)	19%	58%		62%		20%	55%		62%	
Wei	1,266	44%	2.6%	15%	(-)	30%	53%		54%		30%	54%		54%	
Unw	379	83%	0.0%	n.a.		5%	17%	(-)	85%	(+)	4%	0%	(-)	85%	(+)
Sep	846	63%	1.3%	10%	(-)	27%	59%		75%	(+)	25%	62%		75%	(+)
Poo	799	65%	1.3%	20%	(-)	22%	43%		46%		22%	39%		47%	
One	639	58%	2.6%	0%	(-)	17%	8%	(-)	49%		15%	9%	(-)	51%	
Ad1	517	66%	0.6%	33%		27%	65%		68%		27%	65%		67%	
Ad2	489	68%	0.8%	50%		31%	71%		69%		32%	65%		69%	
L1	869	43%	0.0%	n.a.		7%	36%		81%	(+)	8%	30%		82%	(+)
L2	250	83%	0.2%	100%	(+)	49%	74%		43%		49%	74%		43%	
Ma	526	65%	3.8%	11%	(-)	41%	44%		37%		39%	45%		36%	
2Kat3	762	50%	0.6%	100%	(+)	20%	45%		58%		19%	47%		58%	
2Kat1	617	59%	0.0%	n.a.		19%	49%		59%		19%	48%		60%	
4Kat2	266	82%	3.4%	0%	(-)	50%	62%		73%		50%	60%		73%	

Table 65: Results of the extreme bounds analysis for the independent variable DivorceAge aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of age and being divorced (captured by the interaction effect DivorceAge) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 1,645 dependent variables reveals a vague tendency for a negative sign of DivorceAge that can be interpreted in two alternative ways: first, as a vague tendency for an additional negative divorce-effect when divorced migrants turn one year older compared to non-divorced migrants; second, as a vague tendency for an additional negative age-effect for divorced migrants compared to non-divorced migrants.

Note that the 64% of the 4,536 dependent variables that exhibit quasi-complete separation do not contribute to the insight of the interaction effect. The reason is that for 99.8% of the 2,891 dependent variables where quasi-complete separation occurs, all divorced migrants are risk-averse irrespective of their age. Therefore, I fully rely on the statistical analysis of the 1,645 dependent variable that do no exhibit quasi-complete separation to judge the relation of DivorceAge on the probability of being risk-averse.

4.3.12.1.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of DivorceAge by Leamer's criterion over the 1,645 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, the fragility of DivorceAge by Sala-I-Martin's criterion over the 1,645 dependent variables statistically investigated is sensitive to whether Head's individual income or family income is considered (denoted by Ind and Fam) in the way that the influence of DivorceAge becomes robust for the subgroup referring to family income (Fam).

Besides this, two findings are noteworthy: First, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to weighted samples (Wei) is about 6 times (7.5 times) the percentage of robust relations in the subgroup referring to unweighted samples (Unw). Second, comparing results of subgroups referring to different transformation rules, I find that when only migrants with least pronounced degree of risk-attitude (4Kat2) are included in the sample, the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria are more than double the percentages in the subgroup referring to the other transformation rules (2Kat3 and 2Kat1).

Tendency for a sign

The finding of a vague tendency for a positive sign of DivorceAge for the 1,645 dependent variables statistically investigated is sensitive to all different ways to estimate risk-attitudes. In detail, DivorceAge is sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.12.2 Interaction of Divorce and Edu2

4.3.12.2.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable DivorceEdu2 captures the interaction effect of Head being divorced and having a high school diploma or an associate's degree as highest education level. For about 67% of the 4,536 dependent variables in my study DivorceEdu2 shows quasi-complete separation.[316] The fact that DivorceEdu2 almost perfectly predicts risk-attitudes for about 67% of the dependent variables should not be overrated since there are no more than only 5 out of 321 migrants in my sample that are divorced and have education level two. Therefore, I focus on the statistical analysis of DivorceEdu2 based on a still sufficiently high sample of 1,492 dependent variables for which no quasi-complete separation occurs.

All statistical results on the influence of DivorceEdu2 on the probability of being risk-averse are summarized in Table 66, p. 333. Based on 1,492 dependent variables, DivorceEdu2 does not exhibit a

[316] Concerning the 3,044 dependent variables (about 67% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 3,014 out of 3,044 all divorced Heads with a education level two (i.e., high school diploma or an associate's degree as highest education level; DivorceEdu2=1) are risk-averse, and (ii) for 30 out of the 3,044 only up to two divorced Heads with education level two (i.e., high school diploma or an associate's degree as highest education level; DivorceEdu2=1) exist, and these Heads are risk-seeking.

robust and therefore significant influence on the probability of being risk-averse by Leamer's criterion, but a robust influence by Sala-I-Martin's weighted and unweighted criteria of robustness. Finally, DivorceEdu2 shows a strong tendency for a negative sign.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	1,492	67%	15%	0%	(-)	99%	11%	(-)	11%	(-)	89%	0%	(-)	6%	(-)
Ind	1,386	64%	17%	0%	(-)	99%	12%	(-)	12%	(-)	88%	0%	(-)	7%	(-)
Fam	106	84%	1%	0%	(-)	99%	0%	(-)	0%	(-)	100%	0%	(-)	0%	(-)
Ann	491	68%	15%	0%	(-)	99%	12%	(-)	11%	(-)	89%	0%	(-)	7%	(-)
Wor	502	67%	15%	0%	(-)	100%	11%	(-)	11%	(-)	89%	0%	(-)	6%	(-)
Lif	499	67%	15%	0%	(-)	99%	11%	(-)	11%	(-)	88%	0%	(-)	6%	(-)
Var	994	56%	18%	0%	(-)	99%	17%	(-)	17%	(-)	83%	0%	(-)	9%	(-)
LP2	498	78%	11%	0%	(-)	100%	0%	(-)	0%	(-)	100%	0%	(-)	0%	(-)
Ed1	338	74%	9%	0%	(-)	99%	8%	(-)	8%	(-)	91%	0%	(-)	4%	(-)
Ed2	232	64%	17%	0%	(-)	100%	12%	(-)	12%	(-)	88%	0%	(-)	6%	(-)
Ed3	231	64%	17%	0%	(-)	99%	13%	(-)	13%	(-)	87%	0%	(-)	7%	(-)
Ed4	230	65%	17%	0%	(-)	99%	13%	(-)	13%	(-)	88%	0%	(-)	7%	(-)
Ed5	232	64%	16%	0%	(-)	99%	12%	(-)	12%	(-)	88%	0%	(-)	6%	(-)
Ed6	229	65%	16%	0%	(-)	100%	12%	(-)	12%	(-)	88%	0%	(-)	7%	(-)
Wei	1,149	49%	20%	0%	(-)	99%	11%	(-)	11%	(-)	89%	0%	(-)	5%	(-)
Unw	343	85%	9%	0%	(-)	100%	10%	(-)	10%	(-)	90%	0%	(-)	10%	(-)
Sep	807	64%	12%	0%	(-)	100%	14%	(-)	14%	(-)	86%	0%	(-)	5%	(-)
Poo	685	70%	17%	0%	(-)	99%	8%	(-)	8%	(-)	92%	0%	(-)	8%	(-)
One	585	61%	9%	0%	(-)	98%	16%	(-)	15%	(-)	84%	0%	(-)	15%	(-)
Ad1	463	69%	18%	0%	(-)	100%	8%	(-)	8%	(-)	92%	0%	(-)	0%	(-)
Ad2	444	71%	17%	0%	(-)	100%	9%	(-)	9%	(-)	91%	0%	(-)	1%	(-)
L1	824	46%	22%	0%	(-)	100%	20%	(-)	20%	(-)	80%	0%	(-)	11%	(-)
L2	250	83%	6%	0%	(-)	100%	0%	(-)	0%	(-)	100%	0%	(-)	0%	(-)
Ma	418	72%	15%	0%	(-)	98%	0%	(-)	0%	(-)	99%	0%	(-)	0%	(-)
2Kat3	699	54%	26%	0%	(-)	100%	13%	(-)	13%	(-)	87%	0%	(-)	8%	(-)
2Kat1	557	63%	16%	0%	(-)	100%	13%	(-)	13%	(-)	87%	0%	(-)	6%	(-)
4Kat2	236	84%	1%	0%	(-)	96%	2%	(-)	2%	(-)	97%	0%	(-)	2%	(-)

Table 66: Results of the extreme bounds analysis for the independent variable DivorceEdu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of being divorced and having a high school diploma or an associate's degree as highest education level (as captured by the interaction effect DivorceEdu2) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 1,492 dependent variables reveals a strong tendency for a negative sign of DivorceEdu2 that can be interpreted in two alternative ways: first, as a strong tendency for an additional negative divorce-effect for divorced migrants having a high school diploma or an associate's degree (Edu2=1) compared to non-divorced migrants and divorced migrants having other education levels (i.e., divorced migrants having less than a high school diploma (Edu1=1) and divorced migrants having a bachelor's degree or higher (Edu3=1)); second, as a strong tendency for an additional negative education-effect of having a high school diploma or an associate's degree (Edu2=1) for divorced migrants compared to non-divorced migrants and divorced migrants having other education levels (i.e., divorced migrants having less than a high school diploma (Edu1=1) and divorced migrants having a bachelor's degree or higher (Edu3=1)).

Note that (i) the robust influence of DivorceEdu2 on the probability of being risk-averse by Sala-I-Martin's criterion and (ii) the strong tendency for a negative sign of DivorceEdu2 derived from the statistical analysis of 1,492 dependent variables are not in line with 99% of the remaining 3,044 dependent variables where quasi-complete separation occurs in the way that all divorced migrants having a high school diploma or an associate's degree as highest education level (DivorceEdu2=1) are risk-averse. Therefore, findings of the statistical analysis must be interpreted with caution.

4.3.12.2.2 Sensitivity for subgroups of dependent variables

Robustness

First, the fragility and non-significance of DivorceEdu2 by Leamer's criterion over the 1,492 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Second, the robustness of DivorceEdu2 by Sala-I-Martin's criterion over the 1,492 dependent variables statistically investigated is also not sensitive to different ways to estimate risk-attitudes.

Tendency for a sign

The finding of a strong tendency for a negative sign of DivorceEdu2 for the 1,492 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. In detail, DivorceEdu2 is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), (iii) different risk-measures (denoted by Var and LP2), (iv) different education definitions (denoted by Ed1 to Ed6), (v) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (vi) different types of clustering (denoted by Sep and Poo), (vii) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (viii) different measurements of predictive errors (denoted by L1, L2, and Ma), and not sensitive to (ix) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.12.3 Interaction of Divorce and Edu3

4.3.12.3.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable DivorceEdu3 captures the interaction effect of Head being divorced and having a bachelor's degree or higher as highest education level. Unfortunately, a vast majority of 99% of the 4,536 dependent variables of my study shows quasi-complete separation for DivorceEdu3. The remaining 66 dependent variables (about 1% of 4,536) that do not exhibit quasi-complete separation constitute a sample that is too small to be statistically investigated. Consequently, I do not discuss results of the statistical analysis.

Concerning the 4,470 dependent variables (about 99% of 4,536) where quasi-complete separation occurs, it is noteworthy that (i) for 4,446 out of 4,470 all divorced Heads with a bachelor's degree or higher (DivorceEdu3=1) are risk-averse, and (ii) for 24 out of 4,470 only one divorced Head with a bachelor's degree or higher (DivorceEdu3=1) exists and this Heads is risk-seeking. The fact that DivorceEdu3 almost perfectly predicts risk-attitudes for about 99% of the dependent variables should not be overrated since there are no more than only 7 out of 321 migrants in my sample that are divorced and have a bachelor's degree or higher (DivorceEdu3=1).

4.3.12.3.2 Sensitivity for subgroups of dependent variables

Since dividing a sample of 66 dependent variables further into subgroups to run a sensitivity analysis results in sample sizes too small to be statistically investigated, I do no run a sensitivity analysis for the independent variable DivorceEdu3.

4.3.13 Interactions of age and education and their relation to risk-attitudes

On the one hand, I found that older migrants do not tend to possess a higher or lower probability of being risk-averse than younger migrants (see Part C, Section 4.3.3). On the other hand, I found that migrants with higher levels of education tend to possess a lower probability of being risk-averse compared to migrants having lower levels of education (see Part C, Section 4.3.4). In this section I investigate whether age- and education-effects depend on one another.

4.3.13.1 Interaction of Age and Edu2

4.3.13.1.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable AgeEdu2 captures the interaction effect of Head's age and having a high school diploma or an associate's degree as highest education level. The influence of AgeEdu2 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,485 dependent variables since quasi-complete separation occurs only for about 1% of the 4,536 dependent variables.[317] Consequently, I focus on the statistical analysis of AgeEdu2.

All statistical results on the influence of AgeEdu2 on the probability of being risk-averse are summarized in Table 67, p. 337. Based on the sample of 4,485 dependent variables, AgeEdu2 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, AgeEdu2 is also not significantly related to the probability of being risk-averse. Yet, I observe a vague tendency for a negative sign of AgeEdu2.

[317] Concerning the 51 dependent variables (about 1% of 4,536) where quasi-complete separation occurs, it is noteworthy that all migrants with an education level other than education level two are risk-averse (AgeEdu2=0) and among those with education level two (AgeEdu2=1) there are no more than 9 that are risk-seeking.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,485	1%	0.2%	38%		26%	40%		45%		18%	39%		40%	
Ind	3,858	1%	0.1%	100%	(+)	25%	41%		45%		16%	41%		41%	
Fam	627	3%	0.8%	0%	(-)	35%	34%		42%		27%	33%		38%	
Ann	1,495	1%	0.3%	60%		26%	38%		45%		17%	36%		41%	
Wor	1,495	1%	0.1%	0%	(-)	27%	40%		44%		18%	39%		40%	
Lif	1,495	1%	0.1%	0%	(-)	26%	40%		45%		18%	42%		40%	
Var	2,268	0%	0.1%	100%	(+)	20%	83%	(+)	70%		15%	77%	(+)	62%	
LP2	2,217	2%	0.2%	0%	(-)	33%	13%	(-)	19%	(-)	20%	11%	(-)	18%	(-)
Ed1	1,272	2%	0.5%	17%	(-)	29%	35%		43%		21%	32%		39%	
Ed2	645	0%	0.0%	n.a.		22%	19%	(-)	40%		14%	14%	(-)	35%	
Ed3	636	2%	0.3%	100%	(+)	22%	60%		50%		15%	63%		48%	
Ed4	636	2%	0.0%	n.a.		25%	63%		53%		17%	69%		51%	
Ed5	648	0%	0.0%	n.a.		30%	40%		44%		18%	39%		38%	
Ed6	648	0%	0.0%	n.a.		27%	29%		40%		18%	28%		33%	
Wei	2,229	2%	0.0%	n.a.		21%	53%		44%		13%	49%		38%	
Unw	2,256	1%	0.3%	29%		31%	31%		45%		23%	34%		42%	
Sep	2,256	1%	0.4%	38%		27%	46%		48%		18%	43%		44%	
Poo	2,229	2%	0.0%	n.a.		25%	33%		41%		17%	35%		37%	
One	1,512	0%	0.0%	n.a.		10%	46%		37%		8%	35%		27%	
Ad1	1,491	1%	0.3%	50%		36%	41%		51%		24%	43%		49%	
Ad2	1,482	2%	0.3%	25%	(-)	33%	36%		46%		21%	36%		45%	
L1	1,509	0%	0.1%	100%	(+)	17%	85%	(+)	59%		13%	86%	(+)	56%	
L2	1,506	0%	0.0%	n.a.		26%	30%		43%		18%	23%	(-)	38%	
Ma	1,470	3%	0.4%	17%	(-)	37%	25%	(-)	31%		22%	24%	(-)	27%	
2Kat3	1,512	0%	0.0%	n.a.		22%	12%	(-)	37%		14%	4%	(-)	31%	
2Kat1	1,500	1%	0.0%	n.a.		22%	24%	(-)	39%		13%	14%	(-)	35%	
4Kat2	1,473	3%	0.5%	38%		35%	67%		58%		26%	72%		56%	

Table 67: Results of the extreme bounds analysis for the independent variable AgeEdu2 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of age and having a high school or an associate's degree as highest education level (as captured by the interaction effect AgeEdu2) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,485 dependent variables reveals a vague tendency for a negative sign of AgeEdu2 which can be interpreted in two alternative ways: first, as a vague tendency for an additional negative age-effect when migrants having a high school diploma or an associate's degree (Edu2=1) turn one year older compared to migrants having other education levels (i.e., migrants having less than a high school diploma (Edu1=1) and migrants having a bachelor's degree or higher (Edu3=1)); second, as a vague tendency for an additional negative education-effect of having a high school diploma or an associate's degree (Edu2=1) when migrants turn one year older compared to migrants having other education levels (i.e., migrants having less than a high school diploma (Edu1=1) and migrants having a bachelor's degree or higher (Edu3=1)).

4.3.13.1.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of AgeEdu2 in explaining the probability of being risk-averse over the 4,485 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, four findings are noteworthy: First, comparing results of subgroups referring to Head's individual versus family income, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to family income (Fam) is about 1.4 times (1.7 times) the percentage of robust relations in the subgroup referring to Head's individual income (Ind). Second, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 1.7 times (1.3 times) the percentage of robust relations in the subgroup referring to variance (Var). Third, comparing the results of the subgroups referring to weighted versus unweighted samples, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to unweighted samples (Unw) is about 1.5 times (1.8 times) the percentage of robust relations in the subgroup referring to weighted samples (Wei). Fourth, comparing results of subgroups referring to minimizing different predictive errors, I find that when greater weight is put on greater predictive errors step by step (L1, L2, Ma), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

Tendency for a sign

The finding of a vague tendency for a negative sign of AgeEdu2 for the 4,485 dependent variables statistically investigated is not sensitive to (i) whether Head's individual income or family income is considered (denoted by Ind and Fam), (ii) variations of the planning period (denoted by Ann, Wor, and Lif), and not sensitive to (iii) different types of clustering (denoted by Sep and Poo).

However, the finding of a vague tendency for a negative sign of AgeEdu2 is sensitive to (i) variations of the different risk-measures (denoted by Var and LP2), (ii) different education definitions (denoted by Ed1 to Ed6), (iii) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (iv) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (v) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (vi) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.13.2 Interaction of Age and Edu3

4.3.13.2.1 Results irrespective of the way risk-attitudes are estimated

Significance, robustness, and tendency for a sign

The dummy variable AgeEdu3 captures the interaction effect of Head's age and having a bachelor's degree or higher as highest education level. The influence of AgeEdu3 on the probability of being risk-averse can be statistically investigated for a sufficiently high sample of 4,485 dependent variables since quasi-complete separation occurs only for about 1% of the 4,536 dependent variables.[318] Consequently, I focus on the statistical analysis of AgeEdu3.

All statistical results on the influence of AgeEdu3 on the probability of being risk-averse are summarized in Table 68, p. 340. Based on the sample of 4,485 dependent variables, AgeEdu3 does not exhibit a robust influence on the probability of being risk-averse by any criterion of robustness. Hence, AgeEdu3 is also not significantly related to the probability of being risk-averse. Yet, I observe a strong tendency for a positive sign of AgeEdu3.

---

[318] Concerning the 51 dependent variables (about 1% of 4,536) where quasi-complete separation occurs, it is noteworthy that all migrants with education level three (AgeEdu2<>0) are risk-averse irrespective of their age.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Group of ways to estimate risk-attitude	number of dependent variables	missing values	Leamer			Sala-I-Martin weighted					Sala-I-Martin unweighted				
			significant	positive		CDF(0) ≥95%	positive		positive		CDF(0) ≥95%	positive		positive	
	n	% of 4,536	% of n	% of significant		% of n	% of robust		% of n		% of n	% of robust		% of n	
Total	4,485	1%	0.1%	0%	(-)	23%	75%	(+)	64%		20%	77%	(+)	66%	
Ind	3,858	1%	0.1%	0%	(-)	22%	73%		64%		19%	76%	(+)	66%	
Fam	627	3%	0.0%	n.a.		33%	82%	(+)	69%		26%	81%	(+)	69%	
Ann	1,495	1%	0.2%	0%	(-)	22%	72%		63%		19%	74%		66%	
Wor	1,495	1%	0.0%	n.a.		24%	76%	(+)	65%		20%	77%	(+)	67%	
Lif	1,495	1%	0.0%	n.a.		23%	76%	(+)	64%		20%	78%	(+)	66%	
Var	2,268	0%	0.1%	0%	(-)	15%	61%		56%		11%	55%		61%	
LP2	2,217	2%	0.0%	n.a.		32%	82%	(+)	73%		28%	85%	(+)	72%	
Ed1	1,272	2%	0.0%	n.a.		27%	83%	(+)	67%		22%	84%	(+)	68%	
Ed2	645	0%	0.2%	0%	(-)	20%	70%		68%		19%	76%	(+)	67%	
Ed3	636	2%	0.2%	0%	(-)	18%	69%		59%		14%	64%		61%	
Ed4	636	2%	0.2%	0%	(-)	18%	70%		61%		14%	66%		63%	
Ed5	648	0%	0.0%	n.a.		27%	77%	(+)	64%		23%	83%	(+)	71%	
Ed6	648	0%	0.0%	n.a.		26%	66%		65%		22%	71%		68%	
Wei	2,229	2%	0.0%	n.a.		20%	85%	(+)	74%		16%	86%	(+)	77%	(+)
Unw	2,256	1%	0.1%	0%	(-)	26%	67%		54%		23%	70%		56%	
Sep	2,256	1%	0.0%	n.a.		24%	71%		65%		17%	72%		67%	
Poo	2,229	2%	0.1%	0%	(-)	22%	80%	(+)	63%		22%	80%	(+)	66%	
One	1,512	0%	0.0%	n.a.		10%	17%	(-)	55%		8%	19%	(-)	64%	
Ad1	1,491	1%	0.0%	n.a.		32%	86%	(+)	72%		27%	87%	(+)	69%	
Ad2	1,482	2%	0.2%	0%	(-)	28%	84%	(+)	67%		24%	84%	(+)	66%	
L1	1,509	0%	0.0%	n.a.		13%	79%	(+)	64%		11%	83%	(+)	64%	
L2	1,506	0%	0.0%	n.a.		25%	84%	(+)	64%		19%	87%	(+)	66%	
Ma	1,470	3%	0.2%	0%	(-)	32%	67%		65%		29%	67%		69%	
2Kat3	1,512	0%	0.0%	n.a.		20%	98%	(+)	73%		16%	98%	(+)	77%	(+)
2Kat1	1,500	1%	0.0%	n.a.		21%	92%	(+)	73%		19%	94%	(+)	78%	(+)
4Kat2	1,473	3%	0.2%	0%	(-)	29%	46%		47%		24%	48%		44%	

Table 68: Results of the extreme bounds analysis for the independent variable AgeEdu3 aggregated over all 4,536 dependent variables and subgroups of dependent variables. Where n.a. denotes not available due to a division by zero, (+) denotes a 75% majority for a positive coefficient, (-) denotes a 75% majority for a negative coefficient.

Comparison to the literature and interpretation

Since the interaction of age and having a bachelor's degree or higher as highest education level (as captured by the interaction effect AgeEdu3) has not been surveyed yet, my findings are new to the literature.

The statistical analysis based on 4,485 dependent variables reveals a strong tendency for a positive sign of AgeEdu3 which can be interpreted in two alternative ways: first, as a strong tendency for an additional positive age-effect when migrants having a bachelor's degree or higher (Edu3=1) turn one year older compared to migrants having other education levels (i.e., migrants having less than a high school diploma (Edu1=1) and migrants having a high school diploma or an associate's degree (Edu2=1)); second, as a strong tendency for an additional positive education-effect of having a bachelor's degree or higher (Edu3=1) when migrants turn one year older compared to migrants having other education levels (i.e., migrants having less than a high school diploma (Edu1=1) and migrants having a high school diploma or an associate's degree (Edu2=1)).

4.3.13.2.2 Sensitivity for subgroups of dependent variables

Robustness

The fragility and non-significance of AgeEdu3 in explaining the probability of being risk-averse over the 4,485 dependent variables statistically investigated is not sensitive to different ways to estimate risk-attitudes. Yet, three findings are noteworthy: First, comparing results of subgroups referring to Head's individual versus family income, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to family income (Fam) is about 1.5 times (1.6 times) the percentage of robust relations in the subgroup referring to Head's individual income (Ind). Second, comparing the results of the subgroups referring to different risk-measures, I find that the percentage of robust relations by Sala-I-Martin's weighted (unweighted) criterion in the subgroup referring to semi-variance (LP2) is about 2.1 times (2.5 times) the percentage of robust relations in the subgroup referring to variance (Var). Third, comparing results of subgroups referring to minimizing different predictive errors, I find that when greater weight is put on greater predictive errors step by step (L1, L2, Ma), the percentages of robust relations by Sala-I-Martin's weighted and unweighted criteria increase visibly.

<u>Tendency for a sign</u>

The finding of a strong tendency for a positive sign of AgeEdu3 for the 4,485 dependent variables statistically investigated is not sensitive to whether Head's individual income or family income is considered (denoted by Ind and Fam).

However, the finding of a strong tendency for a positive sign of AgeEdu3 is sensitive to (i) variations of the planning period (denoted by Ann, Wor, and Lif), (ii) variations of the different risk-measures (denoted by Var and LP2), (iii) different education definitions (denoted by Ed1 to Ed6), (iv) variations in the weighting of the sample from which income parameters are estimated (denoted by Wei and Unw), (v)different types of clustering (denoted by Sep and Poo), (vi) whether income parameters are estimated from annual income data of one year or three years of data (denoted by One, Ad1, and Ad2), (vii) different measurements of predictive errors (denoted by L1, L2, and Ma), and sensitive to (viii) different transformation rules (denoted by 2Kat3, 2Kat1, and 4Kat2).

4.3.14 Summary on sensitivity for subgroups of dependent variables

4.3.14.1 Motivation and objective: Which competing solutions are necessary?

In order to derive a statement irrespective of the way risk-attitudes are estimated, it is necessary to analyze all 4,536 different ways to estimate risk-attitudes since it is not clear which of the competing solutions that cause the 4,536 ways to estimate risk-attitudes is the right one. At the same time keeping all competing solutions and its resulting 4,536 ways to estimate risk-attitudes is extremely time consuming and puts many constraints on the methodology that can be applied to analyze the relation of risk-attitudes and socio-economic characteristics (see Part C, Section 4.2.1 for reasons to apply extreme bounds analysis). This raises the question which of the competing solutions that cause the 4,536 different ways to estimate risk-attitudes does not alter the results and can therefore be neglected in future studies using the same data set.

The sensitivity analyses run so far answers this question for each independent variable separately. That is, whether the effect of an independent variable on the probability of being risk-averse systematically differs for competing solutions. These results can now be summarized to answer the question over all independent variables analyzed where the following methodology is applied.

4.3.14.2 Methodology

Sensitivity of statements concerning (i) robustness and significance by Leamer's criterion, (ii) robustness by Sala-I-Martin's criterion, and (iii) tendency for a sign

To decide whether a certain group of competing solution, e.g., risk-measures, can be neglected in future studies using the same data set, I count the number of independent variables whose statements concerning (i) significance and robustness by Leamer's criterion, (ii) robustness by Sala-I-Martin's criterion, and (iii) tendency for a sign are sensitive to variations of competing solutions. I investigate sensitivity concerning the three statements (i) to (iii) separately since, for example, different risk-measures might change statements on significance but not on the tendency for a sign. This differentiation enables potential future researchers to include competing solutions depending on their research interest, i.e., (i) interest on significance by Leamer's criterion, (ii) interest in robustness by Sala-I-Martin's criterion, and/or (iii) interest in the tendency for a sign. In the following sections summary tables list which independent variables' statements concerning (i) to (iii) change for which competing solutions. The Greek letter delta (Δ) in Table 69, p. 347 denotes whether the statement of a certain independent variable is sensitive to variations of the respective competing solution.

Two groups of competing solutions can be further separated into subgroups and are therefore investigated in more detail as follows: First, for competing solutions to the problem of how to define education, I further examine whether sensitivity can be attributed to alternative answers to one of the two questions that create the six education definitions. These question are: (1) The question on how to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File, and (2) whether education variables from years before or after the move should be considered first when education variables are not available in the year of the move. If the sensitivity can be attributed to alternative answers to one of the two questions this is indicated by the symbol X in Table 69, p. 347. If no sensitivity can be observed for an independent variable or the sensitivity cannot be attributed to alternative answers to a certain question, nothing is reported. Second, for competing solutions to different time periods from which income parameters are estimated, I additionally analyze whether the sensitivity can be attributed to either (1) the length of the time period (one year's versus three year's data) or (2) whether inflation adjustment is performed for three year's data. Here again, the symbol X in Table 69, p. 347 indicates whether the sensitivity can be attributed to (1) or (2). If no sensitivity can be observed for an independent variable or the sensitivity cannot be attributed to alternative answers to either (1) or (2), nothing is reported.

<u>Three priority-categories of competing solutions</u>

To simplify the decision of which competing solutions should be included in future analysis using the same data set, I define competing solutions into priority-categories A, B, and C. Priority-category A includes competing solutions that should be included in any analysis since they most certainly change the respective statements concerning (i) robustness and significance by Leamer's criterion, (ii) robustness by Sala-I-Martin's criterion, or (iii) tendency for a sign, respectively. Competing solution of priority-categories B and C are optional to be included, while competing solutions of priority-category B should be given preference to be included over priority category C.

4.3.14.3 Which competing solutions can be neglected

Of the 32 independent variables in my study I ran no sensitivity analysis on DivorceEdu3 because (quasi-) complete separation reduces the number of dependent variables in several subgroups to zero. This makes a sensible sensitivity analysis for DivorceEdu3 impossible. Consequently, the following results relate only to the remaining 31 independent variables of my study.

<u>Sensitivity concerning robustness and significance by Leamer's criterion</u>

Competing solutions do never alter the results of an independent variable by Leamer's criterion of robustness since all independent variables are not robustly and therefore not significantly related to risk-attitudes by Leamer's criterion in any subgroup. Consequently, concerning robustness and significance by Leamer's criterion, I define all competing solutions to belong to priority-category C. This means, a potential future researcher that is only interested in statements based on Leamer's criterion can choose one out of 4,536 ways to estimate risk-attitudes without the danger of misleading results.

<u>Sensitivity concerning robustness by Sala-I-Martin's criterion</u>

Of the 31 independent variables for which a sensitivity analysis is run, the only independent variable whose statement on robustness by Sala-I-Martin's criterion is altered by competing solutions is Divorce. Therefore, I define competing solutions that do not alter the respective statement of any independent variable to belong to priority-category C. This relates to competing solutions on (i) the planning period and (ii) the time period from which income parameters are estimated. In contrast, I define all competing solutions that alter the statement of robustness of Divorce by Sala-I-Martin's criterion to belong to priority-category B. This relates to competing solutions on (i) income that is considered (Head's versus family income), (ii) risk-measure, (iii) education definition, (iv) weighting of the sample, (v) type of clustering, (vi) weighting of predictive errors, and (vii) the transformation rule.

Consequently, a potential future researcher that is only interested in statements based on Sala-I-Martin's weighted and unweighted criteria can neglect variations of competing solutions of priority-category C without the danger of obtaining misleading results. For competing solutions of priority-category B it is optional to include them in the analysis. Since only the statement of the independent variable Divorce is affected by competing solutions of this category, I recommend accounting for competing solutions of priority-category B only if the independent variable Divorce is of interest to the researcher.

Sensitivity concerning the tendency for a sign

In contrast to statements concerning significance and robustness, the tendency for a sign exhibits great sensitivity. Table 69, p. 347 lists for each competing solution and independent variable whether the tendency for a sign is sensitive (denoted by Greek letter delta Δ). It shows that the number of independent variables whose tendency for a sign changes is highest for competing solutions relating to different (i) risk-measures, (ii) weighting of predictive errors, and (iii) transformation rules. Therefore, I define these competing solutions to belong to priority-category A that should not be left out when potential future researchers are interested in the sign of the independent variables.

As competing solutions of the next highest priority-category B, I defined all competing solutions that relate to variations in the way data input is gained, i.e., variations of (i) the education definition, (ii) weighing of the sample, (iii) type of clustering, and (iv) time period from which income parameters are estimated. For these competing solutions it is still a considerable number of independent variables whose tendency for a sign changes, i.e., of the 31 independent variables 10 to 12 change their tendency for a sign for competing solutions of priority-category B. Therefore, I recommend not leaving out competing solutions of this priority-category. Besides this general statement, two further statements can be made. First, concerning competing definitions of education, it is noteworthy that the sensitivity of the 12 independent variables cannot be clearly attributed to competing answers to one of the two questions. The sensitivity of 5 out of 12 independent variables can be attributed to competing answers to Question (1), but there are still 7 out of 12 independent variables whose sensitivity cannot be explained by competing answers on these two questions. Consequently, if potential future researchers decide to account for competing solutions of priority-category B, all six education definitions must be considered. Second, concerning different time periods from which income parameters are estimated, it seems decisive whether income parameters are estimated from one year's data or three year's data, but not whether three year's data is adjusted for inflation or not. Among the 12 independent variables that are sensitive to different time periods from which income parameters are estimated, the sensitivity of 10 independent variables can clearly be attributed to the difference in one year's versus three year's data, while inflation adjustment does not change the results. Consequently, it is only necessary to estimate income parameters from on one year's data versus three year's data, but three year's data does not need to be adjusted for inflation.

The remaining competing solutions I define as priority-category C, namely variations of (i) the income that is considered (Head's versus family income) and (ii) different planning periods. If potential future researchers are interested in the tendency for a sign and want to reduce the number of dependent variables to be analyzed, they can at best leave out competing solutions of priority-category C.

	Male	Age	Edu2	Edu3	AgeSquared	FuMembers	FuMembers	Single	Pair	Family	Divorce	MaleSingle	MalePair	MaleFamily	MaleDivorce	MaleAge	MaleEdu2	MaleEdu3	SingleAge	SingleEdu2	SingleEdu3	PairAge	PairEdu2	PairEdu3	FamilyAge	FamilyEdu2	FamilyEdu3	DivorceAge	DivorceEdu2	AgeEdu2	AgeEdu3	number of Δ	% of Δ	Priority-categories
Model specification																																		
Decision based on Head's versus family income (Ind vs. Fam)	Δ						Δ													Δ												3	10%	C
Planning period (Ann vs. Wor vs. Lif)																				Δ								Δ			Δ	3	10%	C
Risk-measure (Var vs. LP2)	Δ	Δ		Δ	Δ		Δ										Δ	Δ	Δ	Δ		Δ		Δ				Δ		Δ	Δ	14	45%	A
Ways to gain data input																																		
Education definition (Ed1 vs. Ed2 vs. Ed3 vs. Ed4 vs. Ed5 vs. Ed6)	Δ						Δ		Δ				Δ				Δ	Δ	Δ	Δ						Δ		Δ		Δ	Δ	12	39%	B
Question (1): How to treat opposing education information in the Single-Year Family Files and the Cross-Year Individual File (Ed1/Ed2 vs. Ed3/Ed4 vs. Ed5/Ed6)	X						X		X				X																		X	5	16%	
Question (2): Whether education variables from years before or after the move should be considered first (Ed1/Ed3/Ed5 vs. Ed2/Ed4/Ed6)																																0	0%	
Weighting of sample (Wei vs. Unw)				Δ					Δ				Δ				Δ	Δ	Δ	Δ								Δ		Δ	Δ	10	32%	B
Type of clustering (Sep vs. Poo)	Δ						Δ		Δ				Δ				Δ	Δ		Δ						Δ		Δ			Δ	10	32%	B
Time periods from which income parameters are estimated (One vs. Ad1 vs. Ad2)	Δ	Δ					Δ			Δ							Δ	Δ		Δ				Δ	Δ	Δ		Δ			Δ	12	39%	B
One year's data versus three year's data (One vs. Ad1/Ad2)	X	X					X			X								X		X				X	X	X		X			X	11	35%	
Three year's data adjusted for inflation or not (Ad1 vs. Ad2)	X																															1	3%	
Estimation procedure																																		
Weighting of predictive errors (L1 vs. L2 vs. Ma)	Δ								Δ				Δ					Δ	Δ	Δ		Δ		Δ	Δ	Δ	Δ	Δ		Δ	Δ	14	45%	A
Transformation to binary dependent variable																																		
Transformation rules (2Kat3 vs. 2Kat1 vs. 4Kat2)	Δ	Δ		Δ			Δ		Δ	Δ		Δ	Δ				Δ	Δ	Δ	Δ		Δ		Δ	Δ	Δ	Δ	Δ		Δ	Δ	20	65%	A

Table 69: Overview of sensitivity concerning the tendency for a sign due to competing solutions by independent variables.
Where Δ denotes independent variable's whose tendency for a sign is sensitive to competing solutions, A, B, and C denote the priority-categories of competing solutions where competing solutions of priority-category A should not be left out and those of priority-categories B and C might be left out after carefully weighting up time efficiency and well-grounded statistical statements. Note that no sensitivity analysis has been run on DivoreEdu3 since (quasi-) complete separation reduced the number of dependent variables to zero in several subgroups.

5 Conclusion

Research questions and results

My empirical dissertation is concerned with three research questions: First, what are migrants' individual risk-attitudes in the context of economic migration? Second, does risk actually plays a significant role in the economic migration decision, i.e., are economic migrants significantly different from being risk-neutral? Third, how are economic migrants' risk-attitudes related to their socio-economic characteristics such as gender, age, and education?

To isolate the effect of economic risk, U.S. interstate migration is considered. The empirical analysis is based on a mean-variance migration decision model that accounts for all types of risk-attitudes (i.e., risk-averse, risk-neutral, and risk-seeking migrants). Since a suitable data set does not exist, it must be created. This is afflicted with multiple complex problems for which several competing solutions exist - all of which are equally reasonable. Therefore, there is not one single way, but 4,536 ways to estimate risk-attitudes of the 321 economic migrants in my sample.

Concerning the first research question, I find that irrespective of the way risk-attitudes are estimated, the 321 migrants of my sample include both risk-averse and risk-seeking migrants. Consequently, they cannot be considered to be homogeneously risk-neutral or risk-averse.

Concerning the second research question, I find that, first, irrespective of the way risk-attitudes are estimated, risk actually plays a significant role in the economic migration decision. Second, irrespective of the way risk-attitudes are estimated, migrants are on average significantly risk-averse.

Concerning the third research question, I conclude that the relation between socio-economic characteristics and risk-attitudes of economic migrants crucially depends on the way risk-attitudes are estimated. Therefore, a clear-cut statement on significance, sign, and robustness of the 32 socio-economic characteristics irrespective of the way risk-attitudes are estimated cannot be made. This "non-result" is due to the fact that I am interested in the general relation between risk-attitudes and socio-economic characteristics (i.e. irrespective of the way risk-attitudes are estimated). In contrast, the literature analyzes the relation between risk-attitudes and socio-economic characteristics within one particular model. However, such an approach cannot answer my third research question.

Practical implication 1

Although it was necessary to keep all 4,536 ways to estimate risk-attitudes in my dissertation, this approach is extremely time consuming and puts many constraints on the methodology that can be applied. Consequently, I tried to identify competing solutions that can be neglected by potential future researchers using the same data set since the competing solutions do not alter the results of my study. Running far ranging sensitivity analyses on competing solutions, I find that, first, if the researcher is interested in statements according to Leamer’s criterion, he can choose one reasonable way out of 4,536 to estimate risk-attitudes. Second, if the researcher is interested in statements according to Sala-I-Martin’s weighted/unweighted criterion and is not investigating the relation of Divorce, he can again choose one out of 4,536 ways to estimate risk-attitudes. Third, when the researcher is interested in the tendency for a sign, he must carefully weigh up between time efficiency and statistical statements based on a great variety of different ways to estimate risk-attitudes.

If a potential future researcher uses another data set than the one used for my dissertation, he is left with only 36 out of 4,536 different ways to estimate risk-attitudes. The remaining 4,500 of the 4,536 ways to estimate risk-attitudes result from the structure of the specific data sets used for my study. Consequently, potential future researchers using another data set only need to consider (i) Head’s versus family income, (ii) different planning periods, (iii) different risk-measures, and (iv) different weighting of predictive errors.

Practical implication 2

To derive policy implications on how the desired group of young and well-educated migrants can be attracted, it is necessary to know how socio-economic characteristics relate to risk-attitudes. My dissertation shows that this relation crucially depends on the way risk-attitudes are estimated. This means that first, policy implications based on only one single or only a few ways to estimate risk-attitudes are misleading. Second, it is not possible to attract a specific group of migrants by introducing a self-selecting risk policy. An example for such a risk policy is offering migrants well-paid jobs without offering access to social security payments. This risk policy rests upon the assumption that only the desired group of young and well-educated migrants is risk-seeking and will therefore be attracted by this specific risk policy. However, since there is no strict relation between socio-economic characteristics and risk-attitudes, such a risk policy could also attract risk-seeking but less skilled migrants.

References

Alexander, J. Trent, Michael Davern, and Betsey Stevenson. 2010. "Inaccurate Age and Sex Data in the Census PUMS Files: Evidence and Implications." CESifo Working Paper No. 2929.

Anam, Mahmudul, Shin-Hwan Chiang, and Lieng Hua. 2004. "Uncertainty and International Migration: An Option cum Portfolio Model." *Journal of Labor Research,* 29(3): 236-250.

Andreski, Patricia, Katherine McGonagle, and Robert Schoeni. 2007. "An Analysis of the Quality of the Health Data in the Panel Study of Income Dynamics." Panel Study of Income Dynamics Technical Paper Series #09-02. Michigan: Institute for Social Research, Survey Research Center, University of Michigan, Ann Arbor, MI.
http://psidonline.isr.umich.edu/Publications/Papers/tsp/2009-02_Quality_Health_Data_PSID_.pdf (accessed May 16, 2011).

Arias, Elizabeth. 2002. „United States life tables, 2000." National vital statistics reports, 51(3). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth. 2004a. „United States life tables, 2001." National vital statistics reports, 52(14). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth. 2004b. „United States life tables, 2002." National vital statistics reports, 53(6). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth. 2006. „United States life tables, 2003." National vital statistics reports, 54(14). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth. 2007. „United States life tables, 2004." National vital statistics reports, 56(9). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth. 2010. „United States life tables, 2006." National vital statistics reports, 58(21). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth. 2011. „United States life tables, 2007." National vital statistics reports, 59(9). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Arias, Elizabeth, Brian L. Rostron, and Betzaida Tejada-Vera. 2010. „United States life tables, 2005." National vital statistics reports, 58(10). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Association of American Medical Colleges. 2012. Admission Requirements.
https://www.aamc.org/students/applying/requirements/ (accessed April 20, 2012).

Backhaus, Klaus, Bernd Erichson, Wulff Plinke, and Rolf Weiber. 2008. *Multivariate Analysemethoden – Eine anwendungsorientierte Einführung*. 12th Ed., Berlin: Springer.

Badunenko, Oleg, Nataliya Barasinska, and Dorothea Schäfer. 2009. "Risk-attitudes and Investment Decisions across European Countries – Are Women More Conservative Investors than Men?" Deutsches Institut für Wirtschaftsforschung Berlin, Discussion Paper 928.

Bank for International Settlement. 2005. "Zero-coupon yield curves: technical documentation." BIS Papers No 25.
http://www.bis.org/publ/bppdf/bispap25.pdf (accessed July 12, 2012).

Barber, Brad M., and Terrance Odean. 2001. "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics*, 116 (1): 261-292.

Barsky, Robert B., F. Thomas Juster, Miles S. Komball, and Matthew D. Shapiro. 1997. Reference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics*, 112(2): 537-579.

Bauer, Thomas K. 2002. „Migration, Sozialstaat und Zuwanderungspolitik." IZA Discussion Paper No. 505.

Bayer, Christian, and Falko Juessen. 2008. "On the Dynamics of Interstate Migration: Migration Costs and Self-Selection." IZA Discussion Paper No. 3330.

Bertocchi, Graziella, Marianna Brunetti, and Costanza Torricelli. 2008. "Portfolio Choices, Gender and Marital Status." *Rivista di Politica Economica, SIPI Spa*, 98(5): 119-154.

Blais, Ann-Renée, and Elke U. Weber. 2006. „A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations." *Judgment and Decision Making*, 1(1): 33-47.

Bonin, Holger, Amelie Constant, Konstantinos Tatsiramos, and Klaus F. Zimmermann.2006. "Native-Migrant Differences in Risk-attitudes." IZA Discussion Paper No. 1999.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review,* 77(4): 531-553.

Borjas, George J. 1994. "The Economics of Immigration." *Journal of Economic Literature*, XXXII: 1667-1717.

Borjas, George J. 1999a. "Immigration and Welfare Magnets." *Journal of Labor Economics*, 17(4): 607-637.

Borjas, George J. 1999b. "The Economics Analysis of Immigration." In *Handbook of Labour Economics*, Vol. 3, eds. O. Ashenfelter and D. Card. Amsterdam: Elservier.

Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo. 1992. "Self-Selection and International Migration in the Unites States." NBER Working Paper No. 4002.

Bortz, Jürgen, and Christof Schuster. 2010. *Statistik für Human- und Sozialwissenschaftler*. Springer. http://www.springerlink.com/content/978-3-642-12769-4/#section=858908&page=1&locus=7 (accessed August 8, 2011).

Bosco, Luigi. 2000. "Migration and wage flexibility." Working Paper.

Brücker, Herbert, and Cécily Defoort. 2006. "The (Self-) Selection of International Migrants Reconsidered: Theory and New Evidence." IZA Discussion Paper No. 2052.

Brücker, Herbert, and Parvati Trübswetter. 2004. "Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany." IZA Discussion Paper No. 986.

Burda, Michael C. 1995. “Migration and the Option Value of Waiting.” Institute for International Economic Studies, Mimeo.

Buurman, Margaretha, Josse Delfgaauw, Robert Dur, and Seth van den Bossche. 2012. “Public Sector Employees: Risk Averse and Altruistic?” CESifo Working Paper: Behavioural Economics, No. 3851.

Campos, Julia, Neil R. Ericsson, and David F. Hendry. 2005. “General-to-specific Modeling: An Overview and Selected Bibliography.” Board of Governors of the Federal Reserve System International Finance Discussion Papers Number 838.

Chen, Kong-Pin, Shin-Hwan Chiang, and Siu Fai Leung. 2001. “Migration, Family, and Risk Diversification.“ Working Paper.

Compton, Janice, and Robert. A. Pollak. 2004. *“Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas?”* NBER Working Paper No. 10918.

Constant, Amelie, and Klaus F. Zimmermann. 2005. “Immigrant Performance and Selective Immigration Policy: A European Perspective.” IZA Discussion Paper No. 1715.

Daveri, Francesco, and Riccardo Faini. 1999. “Where do migrants go?” *Oxford Economic Papers*, 51: 595-622.

Davies, Paul S., Michael J. Greenwood, and Haizheng Li. 2001. “A Conditional Logit Approach to U.S. State-to-State Migration.” *Journal of Regional Science*, 41(2):337-360.

Demiralp, Berna. 2009. „The Impact of Information on Migration Outcomes.” SSRN Working Paper. http://ssrn.com/abstract=1431069 (accessed February 5, 2013).

Dequiedt, Vianney, and Yves Zenou. 2011. “International Migration, Imperfect Information, and Brain Drain.” IZA Discussion Paper No. 5786.

Deutsche Bundesbank. 1997. “Schätzung von Zinsstrukturkurven.” In *Deutsche Bundesbank Monatsbericht Oktober 1997*, 61-66. Frankfurt: Deutsche Bundesbank. http://www.bundesbank.de/Redaktion/DE/Downloads/Veroeffentlichungen/Monatsberichtsaufsaetze/1997/1997_10_zinsstrukturkurven.pdf?__blob=publicationFile (accessed June 20, 2012).

Ding, Xiaohao, Joop Hartog, and Yuze Sun. 2010. “Can We Measure Individual Risk-attitudes in a Survey?” IZA Discussion Paper No. 4807.

Dohmen, Thomas, Armin J. Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2005. “Individual Risk-attitudes: New Evidence from a Large, Representative, Experimentally - Validated Survey.” IZA Discussion Paper No. 1730.

Dohmen, Thomas, Armin J. Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2009. “Individual Risk-attitudes: Measurement, Determinants, and Behavioral Consequences.” ROA Research Memorandum Series 2009/6.

Donkers, Bas, Bertrand Melenberg, and Arthur van Soest. 2001. "Estimating Risk-attitudes Using Lotteries: A Large Sample Approach." *Journal of Risk and Uncertainty*, 22(2): 165-95.

Dostie, Benoit, and Pierre Thomas Léger. 2006. “Self-Selection in Migration and Returns to Unobservable Skills.” IZA Discussion Paper No. 1942.

Dustmann, Christian. 1996. “An Economic Analysis of Return migration.” Working Paper.

Dustmann, Christian. 2001. “Return Migration, Wage Differentials, and the Optimal Migration Duration.” IZA Discussion Paper No. 264.

d’Haultfoeuille, Xavier, and Arnaud Maurel. 2009. “Inference on a Generalized Roy Model, with an Application to Schooling Decisions in France.” IZA Discussion Paper No. 4606.

Epstein, Gil S. 2002. “Informational Cascades and Decision to Migrate.” IZA Discussion Paper No. 445.

Epstein, Gil S., and Ira N. Gang. 2004. “The Influence of Others on Migration Plans.” IZA Discussion Paper No. 1244.

EUROSTAT, RAMON Eurostat's Metadata Server. 2013. “SCL - International Standard Classification of Occupations for European Union purposes (ISCO-88(COM)).” http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=CL_ISCO88C&StrLanguageCode=EN&IntPcKey=&StrLayoutCode=HIERARCHIC (accessed August 19, 2013).

Everitt, Brian S., Sabine Landau, Morven Leese, and Daniel Stahl. 2011. *Cluster Analysis*. West Sussex, United Kingdom: John Wiley & Sons Ltd. http://books.google.de/books?id=w3bE1kqd-48C&printsec=frontcover&dq=cluster+analysis&hl=de&ei=ZqW6Tqy4BZSDhQfKla2vBw&sa=X&oi=book_result&ct=result&resnum=2&ved=0CDcQ6AEwAQ#v=onepage&q=proximity&f=false (accessed November 8, 2011).

Farley, Reynolds. 1998. *New American Reality - Who We are, How We Got Here, Where We are Going*. New York: Russell Sage Foundation. http://books.google.de/books?id=dKWCMD1V3fIC&pg=PA279&lpg=PA279&dq=%22Mover+versus+migrant%22&source=bl&ots=UYZb1KwAJm&sig=GFibLy3GfjZQMa9wjOCVsv7xMjo&hl=de&ei=oCl1TcLdHNSAhQejr41O&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBgQ6AEwAA#v=onepage&q=%22Mover%20versus%20migrant%22&f=false (accessed January 10, 2013).

Giannetti, Mariassunta. 1999. “On the Mechanics of Migration Decisions: Skill Complementarities and Endogenous Price Differentials.” Banca d’Italia Research Department.

Gibson, John, and David McKenzie. 2009. “The Microeconomic Determinants of Emigration and Return Migration of the Best and Brightest, Evidence from the Pacific.” World Bank Policy Research Working Paper 4965.

Gordon, Peter, and Denis Lawton. 2003. *Dictionary of British Education*. London: Woburn Press.

Grazier, Suzanne, and Peter J. Sloane. 2006. “Accident Risk, Gender, Family Status and Occupational Choice in the UK.” IZA Discussion Paper No. 2302.

Guler, Bulent, Fatih Guvenen, and Giovanni L. Violante. 2010. “Joint-Search Theory: New Opportunities and New Frictions.” Working Paper.

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright. 2006. “The U.S. Treasury Yield Curve: 1961 to the Present.” Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C. http://www.federalreserve.gov/pubs/feds/2006/ (accessed February 16, 2012).

Hallahan, Terrence A., Robert W. Faff, and Michael D. McKenzie. 2004. “An empirical investigation of personal financial risk tolerance”. *Financial Services Review*, 13: 57–78

Halliday, Timothy J. 2008. “Migration, Risk and the Intra-Household Allocation of Labor in El Salvador.” IZA Discussion Paper No. 3322.

Harris, John R., and Michael P. Todaro. 1970. “Migration, Unemployment and Development: A Two-Sector Analysis.” *American Economic Review*, 60(1): 126-142.

Harrison, Glenn W., Morten I. Lau, and E. Elisabet Rutström. 2007. "Estimating Risk-attitudes in Denmark: A Field Experiment.” *Scandinavian Journal of Economics*, 109(2): 341–368.

Hartog, Joop, Ada Ferrer-i-Carbonell, and Nicole Jonker. 2002. “Linking Measured Risk-aversion to Individual Characteristics.” *KYKLOS*, 55(1): 3-26.

Harvard Law School. 2012a. Doctor of Juridical Science (S.J.D.) Program. http://www.law.harvard.edu/prospective/gradprogram/sjd/index.html (accessed April 20, 2012).

Harvard Law School. 2012b. LL.M. Program: Eligibility Requirements and Admissions Criteria. http://www.law.harvard.edu/prospective/gradprogram/llm/eligibility/index.html (accessed April 20, 2012).

Heitmüller, Axel. 2002. “Unemployment benefits, risk-aversion, and migration incentives.” IZA Discussion Paper No. 610.

Heitmüller, Axel. 2005. “Unemployment benefits, risk-aversion, and migration incentives.” *Journal of Population Economics,* 18: 93-112.

Hendry, David F. 1993. *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Holt, Charles A., and Susan K. Laury. 2002. “Risk-aversion and Incentive Effects.” *American Economic Review*, 92(5): 1644-1655.

Institute for Social Research, Survey Research Center, University of Michigan. 1982. “A Panel Study of Income Dynamics, Procedures and Tape Codes, Documentation, 1981 Interviewing Year, Wave XIV, A supplement.” Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed November 26, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 1988. “A Panel Study of Income Dynamics, Procedures and Tape Codes, Documentation, 1985 Interviewing Year, Vol. 1, Wave XVIII, A supplement.” Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed November 26, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 1992a. „A Panel Study of Income Dynamics: Procedures and Tape Codes, 1989 Interviewing Year, Wave XXII, Volume I: Procedures and Tape Codes.“ Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 12, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 1992b. „A Panel Study of Income Dynamics: Procedures and Tape Codes, 1989 Interviewing Year, Wave XXII, Volume II:.“Numerical And Employment Indexes.” Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 12, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 1995a. „A Panel Study of Income Dynamics: Procedures and Tape Codes, 1990 Interviewing Year, Wave XXIII, Volume I: Procedures and Tape Codes.“ Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 12, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 1995b. „A Panel Study of Income Dynamics: Procedures and Tape Codes, 1991 Interviewing Year, Wave XXIV, Volume I: Procedures and Tape Codes.“ Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 12, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 1995c. „A Panel Study of Income Dynamics: Procedures and Tape Codes, 1992 Interviewing Year, Wave XXV, Volume I: Procedures and Tape Codes.“ Ann Arbor, Michigan: Survey Research Center. http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 12, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2007. „ Codebook, Panel Study of Income Dynamics: 2005 Public Release Family File.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed November 26, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008a. „A Panel Study of Income Dynamics: 1994 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 16, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008b. „A Panel Study of Income Dynamics: 1995 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 16, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008c. „A Panel Study of Income Dynamics: 1996 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 16, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008d. „A Panel Study of Income Dynamics: 1997 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 13, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008e. „Panel Study of Income Dynamics: 1999 Public Release Family File, PSID: 1999 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 10, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008f. „Panel Study of Income Dynamics: 2001 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed November 26, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2008g. „1993 Family Data: Final Release Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 16, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2010a. „Panel Study of Income Dynamics: 1968-2007 Individual Data File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed December 13, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2010b. „Panel Study of Income Dynamics: 2003 Public Release Family File Codebook.” Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed November 25, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2010c. „Panel Study of Income Dynamics: 2007 Public Release Family File Codebook." Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed November 25, 2010).

Institute for Social Research, Survey Research Center, University of Michigan. 2010d. Questionnaire of the Panel Study of Income Dynamics, Interview Year 2007. Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/Guide/documents.aspx (accessed July 27, 2011).

Institute for Social Research, Survey Research Center, University of Michigan. 2011a. All Frequently Asked Questions. Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/Guide/FAQ.aspx?Type=ALL (accessed May 10, 2011).

Institute for Social Research, Survey Research Center, University of Michigan. 2011b. "Panel Study of Income Dynamics: 1968-2009 Individual Data File, Codebook." Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed July 29, 2011).

Institute for Social Research, Survey Research Center, University of Michigan. 2011c. „ Panel Study of Income Dynamics: 2009 Public Release Family File Codebook." Ann Arbor, Michigan: Survey Research Center.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (accessed July 27, 2012).

Institute for Social Research, Survey Research Center, University of Michigan. 2011d. "PSID File Structure and Merging PSID Data Files." Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/Guide/FileStructure.pdf (accessed May 10, 201).

Institute for Social Research, Survey Research Center, University of Michigan. 2011e. Sponsors of the Panel Study of Income Dynamics. Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/Guide/FAQ.aspx?Type=ALL (accessed May 10, 201).

Institute for Social Research, Survey Research Center, University of Michigan. [No date a]. "Panel Study of Income Dynamcis." Power-Point Presentation about the Panel Study of Income Dynamics. Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/guide/psidppt/index.ppt (accessed May 11, 201).

Institute for Social Research, Survey Research Center, University of Michigan. [No date b]."PSID Data on Families Since 1968-A National Study of Socioeconomics & Health Over Lifetimes & Across Generations." Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/Guide/Brochures/PSID.pdf (accessed May 10, 201).

Institute for Social Research, Survey Research Center, University of Michigan. [No date c]. Questionnaire of the Panel Study of Income Dynamics, Interview Year 2003. Ann Arbor, Michigan: Survey Research Center.
http://psidonline.isr.umich.edu/Guide/documents.aspx (accessed November 25, 2010).

International Labour Organization. 2004. "Child Labour Survey Data Processing And Storage of Electronic Files: A Practical Guide." Geneva, Switzerland: International Labour Organisation.

Jaeger, David, Holger Bonin, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde. 2007. "Direct Evidence on Risk-attitude and Migration." IZA Discussion Paper No. 2655.

Jaeger, David, Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. 2008. "Direct Evidence on Risk-attitude and Migration." Working Paper.

Janssen, Jürgen, and Wilfried Laatz. 2007. *Statistische Datenanalyse mit SPSS für Windows, eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests*. Springer. http://www.springerlink.com/content/978-3-540-72977-8/#section=348376&page=1 (accessed August 6, 2001).

Jianakoplos, Nancy Ammon, and Alexandra Bernasek. 1998. "Are Women more risk averse?" *Economic Inquiry*, 36(4): 620-630.

Johnson, Joseph G., Andreas Wilke, and Elke U. Weber. 2004. "Beyond a trait view of risk taking: A domain-specific scale measuring risk perceptions, expected benefits, perceived-risk-attitudes in German-speaking population." *Polish Psychological Bulletin*, 35(3): 153-163.

Kan, Kamhon. 2003. "Residential mobility and job changes under uncertainty." *Journal of Urban Economics,* 54: 566–586.

Katz, Eliakim, and Oded Stark. 1986. "Labor Migration and Risk-aversion in Less Developed Countries." *Journal of Labor Economics*, 4(1):134-149.

Kennan, John, and James R. Walker. 2008. "The Effect of Expected Income on Individual Migration Decisions." University of Wisconsin, mimeo.

Khwaja, Yasmeen. 2002. "Should I Stay or Should I Go? Migration under Uncertainty: A Real Options Approach." Working Paper.

Kochanek, Kenneth D., Jiaquan Xu, Sherry L. Murphy, Arialdi M. Miniño, and Hsiang-Ching Kung. 2011. "Deaths: Preliminary data for 2009." National vital statistics reports, 59(4). Hyattsville, Maryland: National Center for Health Statistics. http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Krolzig, Hans-Martin, and David F. Hendry. 2000. "Computer automation of general-to-specific model selection procedures." University of Oxford, Department of Economics, Discussion paper Series ISSN 1471-0498.

Krupka, Douglas J., and Kwame Donaldson. 2008. "Wages, Rents and Heterogeneous Moving Costs." Andrew Young School of Policy Studies Research Paper Series, second version.

Leamer, Edward E. 1983. "Let's Take the Con Out of Economometrics." *American Economic Review*, 75(3): 293-307.

Leamer, Edward E. 1985. "Sensitivity Analysis Would Help." *American Economic Review*, 57(3): 308-313.

Leamer, Edward E., and Herman Leonard. 1983. "Reporting the Fragility of Regression Estimates." *Review of Economics and Statistics*, 65(2): 306-317.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review*, 82(4): 942-963.

Liebig, Thomas, and Alfonso Sousa-Poza. 2004. "Migration, Self-Selection and Income Inequality: An International Analysis." *KYKLOS*, 57(1): 125-146.

Locher, Lilo. 2001. "Testing for the Option Value of Migration." IZA Discussion Paper No. 405.

Lomax, Richard G., and Debboe L. Hahs-Vaughn. 2012. *Statistical Concepts: A Second Course*. 4th ed. New York: Routledge.

Mahmood, Talat, and Klaus Schömann. 2003. "On the Migration Decision of IT-Graduates: A Two-Level Nested Logit Model." Discussion Paper SP II 2003 – 22, Wissenschaftszentrum Berlin.

Matthé, Tom, and Guy De Tré. 2010. "The Bipolar Semantics of Querying Null Values in Regular and Fuzzy Databases Dealing with Inapplicability." In *Information Processing and Management of Uncertainty in Knowledge-Based Systems : 13th International Conference, IPMU 2010, Dortmund, Germany, June/July 2012, Proceedings, Part II,* ed. Eyke Hüllermeier, Rudolf Kruse, and Frank Hoffmann, 137-146. Berlin, Heidelberg: Springer.
http://link.springer.com/chapter/10.1007/978-3-642-14058-7_14 (accessed October 21, 2013)

Mayda, Anna Maria. 2005. "International Migration: A Panel Data Analysis of Economic and Non-Economic Determinants." IZA Discussion Paper No. 1590.

McAleer, Michael, Adrian Rodney Pagan, and Paul A. Volker. 1985. "What Will Take the Con Out of Economometrics?" *American Economic Review*, 75(3): 293-307. Referred to in Leamer, Edward E. 1985. "Sensitivity Analysis Would Help." *American Economic Review*, 57(3): 308-313.

Mincer, Jacob. 1978. "Family Migration Decisions." *Journal of Political Economy*, 86 (5): 749-773.

Miniño, Arialdi M., Sherry L. Murphy, Jiaquan Xu, and Kenneth D. Kochanek. 2011. "Deaths: Final Data for 2008." National Vital Statistics Reports. 59(10). Hyattsville, Maryland: National Center for Health Statistics.
http://www.cdc.gov/nchs/products/nvsr.htm (accessed February 14, 2012).

Minnesota Population Center, University of Minnesota. [No date a].Frequently Asked Questions. http://usa.ipums.org/usa-action/faq (accessed July 29, 2011).

Minnesota Population Center, University of Minnesota. [No date b]. Note on the Standardization of ACS/PRCS Income Variables and Other Dollar Amount Variables.
http://usa.ipums.org/usa/acsincadj.shtml (accessed January 31, 2012).

Minnesota Population Center, University of Minnesota. [No date c]. Online dictionary on variables. http://usa.ipums.org/usa-action/variables/group (accessed December 14, 2011).

Minnesota Population Center, University of Minnesota. [No date d]. Sample Design and Estimation in the American Community Survey (ACS) and the Puerto Rico Community Survey (PRCS).
http://usa.ipums.org/usa/voliii/ACSsamp.shtml(accessed December 6, 2010).

Minnesota Population Center, University of Minnesota. [No date e]. Top coded Variables and Corresponding State Means for Values at and Above the Top code.
http://usa.ipums.org/usa/volii/00topcode.shtml (accessed August 8, 2011).

Minnesota Population Center, University of Minnesota. [No date f]. User's Guide.
http://usa.ipums.org/usa/doc.shtml (accessed July 27, 2011).

Mitze, Timo, and Janina Reinkowski. 2010. "Testing the Neoclassical Migration Model: Overall and Age-Group Specific Results for German Regions." Ruhr Economic Papers No. 226.

Moretto, Michele, and Sergio Vergalli. 2005. "Migration Dynamics." Working paper Fondazione Eni Enrico Mattei Note di Lavoro 108.2005.

Muñoz, Francisco J. Callado, and Natalia Utrero González. 2011. "Determinants of Households' Risk." Working Paper.

National Grammar Schools Association. [No date].Frequently Asked Questions. England.
http://www.ngsa.org.uk/faqs.php (accessed April 19, 2012).

Nelson, C. R., and A. F. Siegel. 1987. “Parsimonious Modeling of Yield Curves.” Journal of Business 60, 473-489. Cited in Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright. 2006. “The U.S. Treasury Yield Curve: 1961 to the Present.” Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C. http://www.federalreserve.gov/pubs/feds/2006/ (accessed February 16, 2012).

Nicholson, Nigel, Emma Soane, Mark Fenton-O'Creevy, and Paul Willman. 2005. „Personality and domain-specific risk taking.” *Journal of Risk Research*, 8(2): 157-176.

O’Donnell, Nuala. 2011. “Analysing the Determinants of Attitudes to Risk and Their Role in Pension and Investment Decisions in Ireland and the UK.” *Central Bank of Ireland, Quarterly Bulletin*, 2: 78-09.

Otrachshenko, Vladimir, and Olga Popova. 2012. “Life (Dis)satisfaction and the Decision to Migrate: Evidence from Central and Eastern Europe.“ Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic, Working Paper ISSN 1211-3298.

Pagan, Adrian Rodney. 1987. “Three Econometric Methodologies: A Critical Appraisal.” *Journal of Economic Surveys*, 1(1): 3—24.

Panel Study of Income Dynamics, public use dataset. 2011. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI (2011). Acknowledgement: The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and the National Science Foundation under award number 1157698.
http://simba.isr.umich.edu/Zips/ZipMain.aspx (last accessed July 29, 2011).

Pedersen, Peder J., Mariola Pytlikova, and Nina Smith. 2004. “Selection or Network Effects? Migration Flows into 27 OECD Countries, 1990-2000.” IZA Discussion Paper No. 1104.

Peridy, Nicolas J. 2006. “Welfare Magnets, Border Effects or Policy Regulations: What Determinants Drive Migration Flows into the EU?” *Global Economy Journal*, 6(4): 1-31.

Powell, Melanie, and David Ansic. 1997. “Gender differences in risk behavior in financial decision-making: An experimental analysis.” *Journal of Economic Psychology*, 18: 605-628.

Princeton University. [No date]. WordNet Online Dictionnary.
http://wordnetweb.princeton.edu/perl/webwn?s=honorary+degree&sub=Search+WordNet&o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&h= (accessed April 20, 2012).

Rasch, Björn, Malte Friese, Wilhelm Hofman, and Ewald Naumann. 2006. *Quantitative Methoden Band 2, Einführung in die Statistik*. Springer.
http://www.springerlink.com/content/978-3-540-33309-8/#section=403172&page=1 (accessed November 4, 2011).

Rogers, Andrei, and Luis J. Castro. 1983. „Regional Migration Differentials In IIASA Nations.“ Working Paper WP 83-40 of the International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rotte, Ralph, and Michael Vogler. 1999. “The Effects of Development on Migration: Theoretical Issues and New Empirical Evidence.” IZA Discussion Paper No. 46.

Roy, A.D. 1951. „Some Thoughts on The Distribution of Earnings.“ *Oxford Economics Papers*, 3: 135-146.

Ruangsiri, Yarika. 2004. “Rural-Urban Migration with Risk-aversion and Regional Uncertainty.” Working Paper. http://www.sed.manchester.ac.uk/research/events/conferences/documents/Arthur%20Lewis%20Papers/Ruangsiri.pdf (accessed July 9, 2013).

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2011. Integrated Public Use Microdata Series of the American Community Survey: Version 5.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2010. Based on the American Community Survey run by the U.S. Census Bureau. http://usa.ipums.org (accessed August 3, 2011).

Säve-Soderbergh, Jenny. 2012. “Self-Directed Pensions: Gender, Risk, and Portfolio Choices.“ *Scandinavian Journal of Economics*, 114(3): 705–728.

Sala-I-Martin, Xavier X. 1997a. “I Just Ran Two Million Regressions.” *American Economic Review*, 87(2): 178-183.

Sala-I-Martin, Xavier X. 1997b. “I Just Ran Four Million Regressions.” NBER Working Paper Series No.6252.

Sarig, Oded, and Arthur Warga. 1989. “Some Empirical Estimates of the Risk Structure of Interest Rates.” *Journal of Finance*, 44 (5): 1351-1360. http://www.jstor.org/stable/info/2328646 (accessed July 12, 2012).

Schich, Sebastian T. 1997. “Schätzung der deutschen Zinsstrukturkurve.” Diskussionspapier 4/97 der Volkswirtschaftlichen Forschungsgruppe der Deutschen Bundesbank. http://www.bundesbank.de/Redaktion/DE/Downloads/Veroeffentlichungen/Diskussionspapiere_1/1997/1997_10_01_dkp_04.pdf?__blob=publicationFile (accessed June 20, 2012).

Sheldon, George. 2002. "Sind die Löhne einkommensmaximierender Personen vergleichbar? Das Roy-Modell im Rückblick". In *Zur Theorie, Empirie und Politik der Einkommensverteilung. Festschrift für Gerold Blümle*, ed. L. Menkhoff and F. Sell, 103-129. Berlin: Springer.

Sinn, Hans-Werner. 2004a. “Migration, Social Standards and Replacement Incomes. How to Protect Low-income Workers in the Industrialized Countries against the Forces of Globalization and Market Integration.” NBER Working Paper No. 10798.

Sinn, Hans-Werner, and Wolfgang Ochel. 2003. “Social Union, Convergence and Migration.” *Journal of Common Market Studies*, 41: 869-896.

Sinn, Hans-Werner, and Martin Werding. 2001. “Immigration following EU Eastern Enlargement.” CESifo Forum.

Sjaastad, Larry A. 1962. “The Costs and Returns of Human Migration.” *Journal of Political Economy*, 70 (5, Part 2): 80-93.

Smith, Terence R. 1979. “Migration, Risk-aversion, and Regional Differentiation.” *Journal of Regional Science*, 19(1): 31-45.

Stark, Oded. 1981. “On the optimal choice of capital intensity in LDCs with migration.” *Journal of Development Economics*, 9: 31-41.

Stark, Oded. 1984. “Discontinuity and the Theory of International Migration.” *KYKLOS*, 37(2): 206-222.

Stark, Oded, and David Levhari. 1982. “On Migration and Risk in LDCs.” *Economic Development and Cultural Change*, 31(1): 191-96.

Stocker, Herbert. 2013. "Kapitel 8, Dummy Variablen." Lecture notes on Empirische Wirtschaftsforschung, University of Innsbruck 2013. http://www.uibk.ac.at/econometrics/einf/kap03_dummy.pdf (accessed October 1, 2013).

Sundén, Annika E., and Brian J. Surette. 1998. "Gender Differences in the Allocation of Assets in Retirement Savings Plans." *American Economic Review*, 88(2): 207-211.

Svensson, Lars E. O. 1994. "Estimating and Interpreting Forward Rates: Sweden 1992-1994." National Bureau of Economic Research Working Paper #4871. http://www.nber.org/papers/w4871.pdf?new_window=1 (accessed June 20, 2012). Cited in Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright. 2006. "The U.S. Treasury Yield Curve: 1961 to the Present." Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C. http://www.federalreserve.gov/pubs/feds/2006/ (accessed February 16, 2012).

Todaro, Michael P. 1969. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries." *American Economic Review*, 59 (1): 138-148.

Umblijs, Janis. 2012. "The effect of networks and risk-attitudes on the dynamics of migration." International Migration Institute Working Paper Series.

United Nations, Department of Economic and Social Affairs, Population Division. 2009. Trends in International Migrant Stock: The 2008 Revision (United Nations database, POP/DB/MIG/Stock/Rev.2008). http://esa.un.org/migration/ (accessed March 27, 2013).

U.S. Bureau of Labor Statistics. 2011. American Community Survey (ACS) Questions and Answers. http://www.bls.gov/lau/acsqa.htm (accessed June 26, 2012).

U.S. Bureau of Labor Statistics. [No date] "CPI Detailed Report Data for August 2011." http://www.bls.gov/cpi/cpid1108.pdf (accessed October 25, 2011).

U. S. Census Bureau. 2003. "American Community Survey, Operation Plan." U.S. Government Printing Office, Washington, DC. http://usa.ipums.org/usa/resources/codebooks/ACS_codebook.pdf (accessed November 3, 2011).

U.S. Census Bureau. 2008. „A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know." U.S. Government Printing Office, Washington, DC. http://www.census.gov/acs/www/Downloads/handbooks/ACSGeneralHandbook.pdf (accessed November 2, 2011).

U.S. Census Bureau. 2009a. „A Compass for Understanding and Using American Community Survey Data: What PUMS Data Users Need to Know." U.S. Government Printing Office, Washington, DC. http://www.census.gov/acs/www/Downloads/handbooks/ACSPUMS.pdf (accessed November 2, 2011).

U.S. Census Bureau. 2009b. „A Compass for Understanding and Using American Community Survey Data: What Researchers Need to Know." U.S. Government Printing Office, Washington, DC. http://www.census.gov/acs/www/Downloads/handbooks/ACSResearch.pdf (accessed November 2, 2011).

U.S. Census Bureau. 2009c. „Design and Methodology. American Community Survey." U.S. Government Printing Office, Washington, DC. http://www.census.gov/acs/www/Downloads/survey_methodology/acs_design_methodology.pdf (accessed November 2, 2011).

U. S. Census Bureau. 2010. Guidance About Income Sources. http://www.census.gov/hhes/www/income/method/guidance/index.html (accessed December 2, 2010).

U.S. Census Bureau. 2012a. About the American Community Survey. http://www.census.gov/acs/www/about_the_survey/american_community_survey/ (accessed June 26, 2012).

U.S. Census Bureau. 2012b. American Community Data Release. http://www.census.gov/acs/www/data_documentation/data_main/ (accessed June 26, 2012).

U.S. Department of Education. 1995. Progress of Education in the United States of America - 1990 through 1994.
http://www2.ed.gov/pubs/Prog95/pt1org.html (accessed April 19, 2012).

U.S. Department of Education, International Affairs Office. 2008a. Structure of the U.S. Education System: Research Doctorate Degrees.
http://find.ed.gov/search?q=cache:mhkiGZcAWzMJ:www2.ed.gov/about/offices/list/ous/international/usnei/us/doctorate.doc+Ph.D.+Doctor&client=default_frontend&output=xml_no_dtd&proxystylesheet=default_frontend&ie=UTF-8&access=p&oe=UTF-8 (accessed April 19, 2012).

U.S. Department of Education, International Affairs Office. 2008b. Structure of the U.S. Education System: First-Professional Degrees.
http://find.ed.gov/search?q=cache:vaGW0zDj8MMJ:www2.ed.gov/about/offices/list/ous/international/usnei/us/professional.doc+doctorate+degrees+US+medicine&client=default_frontend&output=xml_no_dtd&proxystylesheet=default_frontend&ie=UTF-8&access=p&oe=UTF-8 (accessed April 19, 2012).

U.S. Department of the Treasury. 2011a. Interest Rate Statistics. http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/default.aspx (accessed February 16, 2012).

U.S. Department of the Treasury. 2011b. Interest Rates – Frequently Asked Questions. http://www.treasury.gov/resource-center/faqs/Interest-Rates/Pages/faq.aspx (accessed February 16, 2012).

U.S. Social Security Administration. 2011. Retirement benefits by year of birth. http://www.socialsecurity.gov/retire2/agereduction.htm (accessed November 10, 2011).

Vergalli, Sergio. 2006. "The Role of Community in Migration Dynamics." Working paper Fondazione Eni Enrico Mattei Note di Lavoro 4.2006.

Weber, Elke, Ann-Renée Blais, and Nancy E. Betz. 2002. "A Domain Specific Risk-attitude Scale: Measuring Risk Perceptions and Risk Behaviors." *Journal of Behavioral Decision Making*, 15: 263–290.

Wik, Mette, Tewodros Aragie Kebede, Olvar Berglandn and Stein Holden. 2004. „On the Measurement of Risk-aversion from Experimental Data." Department of Economics and Resource Management nor of the Agricultural University of Norway, Discussion Paper #D-16/2004.

Wildasin, David. E. 1991. "Income Redistribution in a Common Labor Market." *American Economic Review*, 81(4): 757 – 774.

Wildasin, David. E. 1994. „Income redistribution and migration." *Canadian Journal of Economics*, 27(3): 637-56.

Wilhelm, Jochen, **and Lars Brüning.** 1992. „Die Fristigkeitsstruktur der Zinssätze Theoretisches Konstrukt und empirische Evaluierung - Untersuchung mit Daten des Kapitalmarktes der Bundesrepublik Deutschland.“ *Kredit und Kapital*, 25: 259-294.

World Bank .2011. Migration and Remittances Factbook 2011. 2nd ed. Washington D.C.: The World Bank.
http://siteresources.worldbank.org/INTLAC/Resources/Factbook2011-Ebook.pdf (accessed February 13, 2013).

Xiao, Jing J., M.J. Alhabeeb, Gong-Soog Hong, and George W. Haynes. 2001. “Attitude toward Risk and Risk-Taking Behavior of Business-Owning Families.” *Journal of Consumer Affairs*, 35(2): 307–325

Yang, Xin-She. 2008. Introduction to Computational Mathematics. Singapore: World Scientific Publishing Co. Pte. Ltd.
http://books.google.de/books?id=T_i0o_55MNwC&pg=PA29&dq=p-norm&hl=de&sa=X&ei=RJrtUNrjGIrNtAa_lYHIDg&ved=0CGcQ6AEwCTgK#v=onepage&q=p-norm&f=false (accessed January 9, 2013).

Yao, Rui, and Sherman D. Hanna. 2005. “The Effect of Gender and Marital Status on Financial Risk Tolerance.” Journal of Personal Finance, 4(1): 66-85.

Appendix

Proof that the preference function used in my dissertation is indeed capable of dealing with risk-averse and risk-seeking individuals

This section proofs exemplary for an exponential utility function where the argument of the function is normally distributed that the preference function of Equation (1) of Part A can be equally applied to risk-averse and risk-seeking decision-makers. In Part A it has simply been argued that choosing α smaller than zero leads to a risk-seeking and α larger than zero to a risk-averse decision-maker.

Expected utility for exponential utility function with normally distributed argument

Start with the expected utility of a decision-maker with exponential utility function

$$E\{U(X)\} = E\left\{-\frac{1}{\alpha}e^{-\alpha X}\right\}.$$

Recall that for a log-normally distributed random variable $Z = e^Y$ (where Y is normally distributed) the expected value reads

$$E(Z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma(Y)} e^{-\frac{1}{2}\cdot\left(\frac{y-E(Y)}{\sigma(Y)}\right)^2} e^y \, dy.$$

To compute the expected value, transform Y into a standard normally distributed random variable $W \equiv \frac{Y-E(Y)}{\sigma(Y)}$ leads to

$$E(Z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2 \cdot \pi}} e^{-\frac{1}{2}\cdot w^2} e^{\mathrm{E}(Y)+\sigma(Y)\cdot w} \, \mathrm{dw},$$

where capital letters denote the random variable and lowercase letters the realizations of the random variable.

Adding $\pm\frac{1}{2} \cdot \sigma^2(Y)$ to the exponent of the above formula leads to

$$E(Z) = \mathrm{e}^{\mathrm{E}(Y)+\frac{1}{2}\cdot\sigma^2(Y)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2 \cdot \pi}} e^{-\frac{1}{2}\cdot\left(w^2-2\cdot\sigma(Y)\cdot w+\sigma^2(Y)\right)} \, \mathrm{dw}$$

and, finally,

$$E(Z) = \mathrm{e}^{\mathrm{E}(Y)+\frac{1}{2}\cdot\sigma^2(Y)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\cdot\pi}} e^{-\frac{1}{2}\cdot(w-\sigma(Y))^2}\,\mathrm{dw}.$$

Using the variable transformation $P \equiv W - \sigma(Y)$ results in

$$E(Z) = \mathrm{e}^{\mathrm{E}(Y)+\frac{1}{2}\cdot\sigma^2(Y)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\cdot\pi}} e^{-\frac{1}{2}\cdot p^2}\,\mathrm{dp}.$$

Since $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\cdot\pi}} e^{-\frac{1}{2}\cdot p^2}\,\mathrm{dp} = 1$, it is eventually obtained

$$E(Z) = \mathrm{e}^{\mathrm{E}(Y)+\frac{1}{2}\cdot\sigma^2(Y)}.$$

Note: In the context of exponential utility it holds $Y \equiv -\alpha X$. Consequently,

$$E(Y) = -\alpha E(X),$$

$$var(Y) = \alpha^2 var(X).$$

Therefore, expected utility can be written as

$$E[U(X)] = -\frac{1}{\alpha} e^{-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)}.$$

<u>**Maximization of preference function for risk-averse decision-makers**</u>

Rational decision-makers will maximize $E[U(X)]$ which means

$$max\left[E[U(X)]\right] = max\left[-\frac{1}{\alpha} e^{-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)}\right].$$

For decision-makers with absolute risk-aversion, it holds $\alpha > 0$ because then the utility function is concave. Since α is a positive constant and the exponential function is larger than zero by definition, maximizing expected utility is equivalent to

$$max\left[-e^{-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)}\right]$$

and, hence,

$$min\left[\underbrace{e^{-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)}}_{>0}\right].$$

Moreover, minimizing an exponential function is equivalent to minimizing its exponent

$$min\left[-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)\right]$$

$$= min\left[-\alpha\left(E(X)-\frac{1}{2}\alpha\cdot var(X)\right)\right].$$

As $-\alpha$ is again smaller than zero for risk-averse decision-makers, this is results in

$$max\left[E(X)-\frac{1}{2}\alpha\cdot var(X)\right].$$

Maximization of preference function for risk-seeking decision-makers

Risk-seeking decision-makers exhibit a negative risk preference parameter ($\alpha < 0$) because then the utility function is convex. Therefore, maximizing expected utility signifies

$$max\left[E[U(X)]\right] = max\left[-\frac{1}{\alpha}e^{-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)}\right]$$

or, since $-\frac{1}{\alpha}$ is positive,

$$max\left[e^{-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)}\right].$$

Moreover, maximizing an exponential function is identical to maximizing its exponent

$$\max\left[-\alpha E(X)+\frac{1}{2}\alpha^2 var(X)\right]$$

$$= \max\left[-\alpha\left(E(X)-\frac{1}{2}\alpha\cdot var(X)\right)\right].$$

As $-\alpha$ is smaller than zero for risk-averse decision-makers, this is equivalent to

$$\max\left[E(X)-\frac{1}{2}\alpha\cdot var(X)\right].$$